The University Of Sheffield.

Clinicians' use of behavioural techniques in CBT: the role of patient and clinician characteristics

By:

Amy Daglish

A thesis submitted in partial fulfilment of the requirements for the award of Doctor of Clinical Psychology at the University of Sheffield

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

Submission Date: May 2018

Declaration

I hereby declare that this work has not been, and will not be, submitted in whole or in part to another University or institution or for the award of any other degree at the University of Sheffield. I confirm that this is my original work and all other sources have been appropriately referenced.

Word Count

Literature Review

Excluding tables, references and appendices	4582
Including tables and references, excluding appendices	7215

Research Report

Excluding tables, references and appendices	7,999
Including tables and references, excluding appendices	11,591

Total

Excluding tables, references and appendices	13,080
Including tables and references, excluding appendices	19,305

Acknowledgments

First of all, I would like to thank my research supervisor, Professor Glenn Waller, for his availability and guidance throughout this project. I would also like to thank my family and friends for their support and encouragement during this journey. Special thanks must go to my sister Emily, for her unfailing belief in me, and to my friend Michelle, for always being on the other end of the phone, and making me take a break to have fun along the way.

Abstract

Evidence-based cognitive behaviour therapy (CBT) protocols vary according to diagnosis but always include two key elements, cognitive change and behavioural change. However, despite the evidence-base and availability of protocols, therapists have been found to 'drift' away from behavioural techniques when delivering CBT. Clinician characteristics, including anxiety and beliefs, have been proposed to contribute to this drift. This project aimed to develop the evidence base by conducting: (i) a systematic review to establish what the current evidence indicates about the relationship between clinicians' beliefs and emotions and their CBT delivery; and (ii) an experimental project evaluating the impact of patient and clinician characteristics on a specific type of exposure, open-weighing, by assessing clinicians' planned weighing behaviours in CBT for eating disorders.

The first part reports a systematic review of 14 studies, assessing the relationship between clinicians' emotions or beliefs and their CBT delivery. Results indicated that clinicians' negative beliefs about exposure and clinicians' anxiety sensitivity lead to less confident exposure delivery, and may be associated with less use of exposure. In addition, other aspects of clinicians' anxiety may be related to less use of behavioural techniques in CBT. These findings support hypotheses around therapist drift that suggest clinicians' beliefs and anxiety lead them to engage in their own safety behaviours, causing them to 'drift' away from evidence-based delivery of behavioural techniques. Limitations of the review are discussed and recommendations for clinical practice and future research are provided.

The second part presents an experimental study assessing the relationship between patient and clinician characteristics and clinicians' planned weighing behaviours. Seventy four clinicians using CBT for eating disorders completed an online survey, reporting their planned weighing behaviours in response to vignettes where

patient characteristics were manipulated. Clinicians also completed measures of clinician characteristics, including firm-empathy, anxiety, beliefs about open weighing, and tendency to make broken leg exceptions (excluding patients from exposure based on patient characteristics, despite a lack of evidence for such exclusions).

Findings indicated that clinicians' negative beliefs about open weighing and tendency to make broken leg exceptions were associated with less guideline-compliant weighing behaviours, and this appeared to be moderated by patient characteristics (diagnosis and weight-related distress). It was concluded that clinicians' beliefs influence their weighing behaviours in CBT for eating disorders, causing them to 'drift' from evidence-based, recommended practice. Limitations of the study are discussed and recommendations for clinical practice and future research are provided.

The two studies contribute to the developing evidence base for therapist drift in CBT. They both suggest that clinicians' negative beliefs about exposure may cause them to avoid exposure or deliver exposure in a less evidence-based manner. Whilst the review suggested that anxiety may also influence clinicians' use of behavioural techniques, this was not found to be the case in the experimental study. Both studies highlight that clinicians using CBT must overcome their negative beliefs to offer an evidence-based approach to patients, which includes confidently delivered behavioural change techniques. Both studies highlight the need for further research to consider observations of real-life clinical practice.

Contents

Access to Thesis Form ii

Declarations iv

Word Counts v

Acknowledgements vi

Abstract vii

Part One: Literature Review 1

Abstract 2

Introduction 4

Method 6

Results 9

Discussion 19

References 25

Appendix A: Adapted Downs and Blacks Critical Appraisal Tool 32

Appendix B: Table Showing Critical Appraisal for Included Studies 39

Part Two: Research Report 41

Abstract 42

Introduction 44

Method 50

Results 57

Discussion 74

References 83

Appendix A: Confirmation of Ethical Approval and Clarification of Title Letter 88

Appendix B: G*Power Output 91

Appendix C: Recruitment Email 92

Appendix D: Intolerance of Uncertainty Scale (IUS-12) 93

Appendix E: Beliefs about Open Weighing Questions 94

Appendix F: Broken Leg Exceptions Scale for Open Weighing (BLES-OW) 95

Appendix G: Firm Empathy Questionaire (FEQ) 96

Appendix H: Core Vignette 97

Appendix I: Weighing Behaviour Questions 98

Appendix J:Participant Information and Informed Consent 99

Appendix K: Participant Debreifing 100

Part One: Literature Review

The impact of clinicians' beliefs and emotions on their CBT delivery

Abstract

Objective: Therapists have been found to 'drift' away from evidence-based behavioural techniques when delivering CBT. Clinicians' anxiety and beliefs have been proposed to contribute to this drift. This systematic review aimed to establish what the current evidence indicates about the relationship between clinicians' beliefs and emotions and their CBT delivery.

Method: A systematic search of three databases (PsycInfo, Medline and Scopus) was conducted. Studies were eligible for inclusion if they considered the relationship between at least one clinician belief or emotion and at least one aspect of CBT delivery. If any exclusion criteria were met, they were excluded. Methodological quality was assessed using an adapted critical appraisal tool which demonstrated good inter-rater reliability in this review. Findings were synthesised in a narrative summary.

Results: 14 papers met eligibility criteria. Of note, anxiety was the only clinician emotion considered. Findings suggest that clinicians' negative beliefs about exposure and clinicians' anxiety sensitivity lead to less confident exposure delivery, and may be associated with less use of exposure. Other aspects of clinicians' anxiety may be related to less use of behavioural techniques in CBT.

Conclusions: These findings support hypotheses around therapist drift that suggest clinicians' beliefs and anxiety lead them to engage in their own safety behaviours and 'drift' away from evidence-based delivery of behavioural techniques. However, further research is needed to replicate and expand findings and overcome methodological flaws, particularly by observing CBT delivery in real-life clinical practice.

Practitioner Points

1. When using CBT, clinicians should be aware that their negative beliefs about exposure and anxiety sensitivity could lead them to deliver exposure more cautiously, or avoid using exposure, despite the strong evidence base for this technique.
2. Clinicians should be prepared to challenge their negative beliefs and anxiety by delivering behavioural elements of CBT, without concurrent use of unsupported adjuncts, and monitoring outcomes.
3. One way to address the 'gap' between the evidence base and clinical practice in CBT may be to address clinicians' beliefs and anxiety in supervision and training. For example this review suggests that increasing clinicians' positive beliefs about exposure could lead to more confident, evidence-based exposure delivery.
4. Further research is needed to overcome methodological flaws in the current evidence base, particularly by observing CBT delivery in real-life clinical practice.

Introduction

Cognitive behavioural therapy (CBT) has been found to be effective in treating a range of mental health difficulties, including depression, eating disorders, and anxiety disorders. This latter category includes post-traumatic stress disorder (PTSD), generalised anxiety disorder (GAD), obsessive compulsive disorder (OCD), social anxiety and panic (e.g., Hay, 2013; Hofmann, Asnaani, Vonk, Sawyer & Fang, 2012; Olatunji, Cisler & Deacon, 2010).

The specific, manualised content of CBT differs according to the disorder being treated, but two key treatment elements are consistently present: cognitive restructuring and behavioural change (e.g., exposure). In light of the evidence for its effectiveness, CBT is recommended by the National Institute for Health and Care Excellence (NICE) as a treatment for several conditions including depression, eating disorders, GAD, panic disorder, PTSD and OCD (NICE, 2005a, 2005b, 2011, 2017, 2018). Many of these recommendations specifically stress the importance of behavioural elements in CBT, such as exposure.

Despite this evidence base, research demonstrates that therapists 'drift' away from evidence-based CBT practice. Clinicians do not always use CBT when it is indicated, or they might report using CBT, but adapt it, making it less likely to be effective (Stobie, Taylor, Quigley, Ewing & Salkovskis, 2007). In addition, research suggests that, whilst clinicians are more likely to use cognitive techniques, they particularly 'drift' away from behavioural techniques, with exposure being particularly under-utilised (e.g., Becker, Zayfert & Anderson, 2004; Becker-Haimes et al., 2017; Hipol & Deacon, 2013; van Minnen, Hendricks & Olff, 2010). When exposure is used, it is often adapted, being delivered cautiously with concurrent use of distress reduction techniques such as relaxation or breathing training (Freiheit, Vye, Swan & Cady, 2004).

These adjuncts to exposure are not evidence-based and reduce the effectiveness of exposure treatment (Deacon et al., 2012; Schmidt et al., 2000).

'Therapist drift' could occur due to a number of factors, including organisational constraints, lack of clinician training, or patient variables. However, it is also possible that clinicians drift due to internal factors, such as their own cognitions, emotions and safety behaviours (Waller, 2009; Waller & Turner, 2016).

Clinicians' poor implementation of exposure-based methods

In exposure, patients maintain contact with their feared stimuli, without using their usual escape or avoidance behaviour, until their anxiety subsides (Foa & McLean, 2016). However, despite its effectiveness, clinicians' use of exposure-based methods is extremely low (e.g., Becker-Haimes et al., 2017).

Clinicians have been found to have a range of negative attitudes to exposure (e.g., Becker et al., 2004; Deacon, Lickel, Farrell, Kemp & Hipol, 2013; Olatunji, Deacon & Abramowitz, 2009; van Minnen et al., 2010). These negative attitudes include beliefs that exposure is unethical, damages the therapeutic relationship and increases risk of drop-out. Some clinicians might believe that exposure is aversive, intolerable and harmful to patients, and that exposure increases the patient's symptoms and risk of decompensation. Some clinicians might also believe that exposure will impact negatively on themselves, through vicarious distress or increasing risk of litigation. In addition, some clinicians have been shown to feel anxious about delivering different CBT elements (Turner, Tatham, Lant & Mountford, 2014). It is possible that these beliefs and anxiety could influence clinicians' delivery of exposure.

Waller and Turner (2016) suggest that clinicians' negative beliefs and anxiety cause clinicians to engage in safety behaviours to alleviate their own concerns. These safety behaviours include avoiding behavioural techniques or reducing the difficulty of such techniques. Waller and Turner argue that clinicians' safety behaviours reinforce the

patients' safety behaviours. Whilst avoiding the behavioural techniques might make the therapist and patient more comfortable in the short-term, this avoidance is likely to be detrimental and maintain or increase anxiety in the longer-term.

Aims

The outstanding question is then, do clinicians' beliefs and anxiety impact on their delivery of CBT, as Waller and Turner suggest? Whilst some narrative reviews have included discussion of some of the literature that may answer this question (e.g., Shafran et al., 2009; Waller, 2016; Waller & Turner, 2016), there has not yet been a systematic literature review to synthesise the empirical research assessing the relationship between these clinician variables and CBT delivery. Therefore, this review will systematically search empirical studies, with an aim of answering the following questions:

1. What is the relationship between clinicians' beliefs and their CBT delivery?
2. What is the relationship between clinicians' emotions and their CBT delivery?

Method

Search Strategy

Three electronic literature databases (PsychInfo, Medline and Scopus) were systematically searched on 15th February 2018. In order to ensure a comprehensive search, synonyms were included in the search terms and these were mapped onto relevant subject headings and 'exploded' to include other related subject headings. No restrictions were applied in terms of date of publication. All literature published up to the date of the search were considered for inclusion.

The following search terms were entered into each database:

("therapist* characteristic*" OR "clinician* characteristic*" OR "therapist* anxiet*" OR "clinician* anxiet*" OR "therapist* mood*" OR "clinician* mood*" OR "therapist* emotion*" OR "clinician* emotion*" OR "therapist* belief*" OR "clinician* belief*"

OR "therapist* attitude*" OR "clinician* attitude*" OR "therapist* factor*" OR "clinician* factor*" OR "therapist* effect*" OR "clinician* effect*")

AND

("therapist* adherence" OR "clinician* adherence" OR "therapy adaptation*" OR "therapist drift" OR "therapist* practice*" OR "clinician* practice*")

AND

("CBT" OR "cognitive* behav*" OR "exposure*")

Inclusion and Exclusion Criteria

Studies were included if they assessed the relationship between at least one clinician emotion or belief and clinicians' use, or delivery style, of at least one CBT technique (e.g., exposure). Use of CBT techniques could include clinicians' decision to use CBT techniques, or their frequency of use of these techniques. Delivery style of CBT techniques could include measures of fidelity, proficiency or concurrent use of adjuncts to CBT techniques (e.g., using controlled breathing alongside exposure). The final inclusion criterion was that papers were published in a peer-reviewed journal. Papers that only assessed clinicians' general attitudes towards evidence-based practice were excluded. Papers were excluded if they were not in English or were not empirical studies (e.g., a narrative review).

Quality Assessment

The methodological quality of the included papers was systematically assessed to aid identification of strengths and limitations. The studies included in this review were diverse in methodology. Therefore, an established quality assurance tool was used, which is appropriate for both randomised and non-randomised studies (Downs & Black, 1998). Some adaptations were made. Item 27 was simplified in accordance with previous studies (e.g., Larson, Vos, & Fernandez, 2013), so that a score of 1 was awarded where studies reported that power was attained, and 0 was awarded where no

sample size calculation was mentioned. Some question wording was also changed to better apply to the papers included in this review (e.g., 'patients' was changed to 'participants'). References to case-control studies were removed, as no papers used this design. See Appendix A for the adapted checklist.

For between-group intervention studies, the total possible score on the critical appraisal tool was 28. However, due to different study designs, not all items were applicable. For example, for cross-sectional survey methodologies, 14 items that related to intervention studies were non-applicable, meaning that the total possible score was 13. In light of the difference in possible scores between studies, quality scores were not categorised using qualitative descriptors (e.g., Larson et al., 2013). Instead, total quality percentage scores were derived by dividing the sum of checklist scores by the number of items assessed, to allow crude comparison between study designs. A higher score was indicative of greater quality, based on the criteria. See Appendix B for the quality assessment grid.

A third year trainee clinical psychologist acted as independent assessor, repeating the quality appraisal on a random 20% (n = 3) of the included studies. Inter-rater reliability was estimated by calculating a two-way mixed-effects intra-class correlation coefficient (ICC) in IBM SPSS version 23, which indicated good inter-rater reliability, ICC = .78 (Koo & Li, 2016). Disagreements in ratings were resolved through discussion. To ensure that the review reflected the breadth of the literature base, no papers were removed based on quality assessment score.

Data Synthesis Method

Statistical analysis was not considered appropriate due to the diversity in methodology, outcomes assessed and measures used across the studies. Therefore, a meta-analysis was not conducted, and data were synthesized in a narrative summary.

Results

A PRISMA diagram (Moher, Liberati, Tetzlaff, Altman, & The PRISMA Group, 2009) is presented in Figure 1 to illustrate the process of selecting papers for inclusion in the review. Database searches yielded 310 results, and 17 duplicate results were removed. The titles and abstracts of the remaining 293 citations were screened against inclusion criteria. Where there was not enough information to determine if papers met the inclusion criteria, they were retained for further full-text review. Twenty papers were deemed relevant for inclusion. The references of these papers were hand searched, which produced a further 11 relevant papers. These 31 full-text articles were assessed against inclusion and exclusion criteria, and 17 were removed. Seven studies did not meet inclusion criteria, and ten met exclusion criteria (two only considered general attitudes to evidence-based practice, one was not available in English, and seven were narrative reviews rather than empirical studies). Fourteen papers were included in the final review, and a summary of key aspects of these papers is included in Table 1.

The studies used diverse methodologies. Nine used cross-sectional, correlational designs, and the majority of these were survey-based. The remaining five studies used quasi-experimental, between-group designs. Of these, four papers used pre-post designs, assessing change following training interventions. The final quasi-experimental design manipulated patient characteristics in vignettes between groups.

Papers assessed a variety of clinician beliefs, but only one clinician emotion was considered - anxiety. Seven papers assessed the impact of clinicians' beliefs, three assessed the impact of anxiety, and four papers considered both clinician beliefs and anxiety. In terms of CBT delivery outcomes, five papers considered the use of a range of CBT techniques and nine focused on use or delivery style of exposure. Papers assessed clinicians working with a range of patient presentations - predominantly anxiety but also other anxiety-based disorders (e.g., panic disorder, OCD, PTSD, eating

disorders). Most papers considered adult client groups, but two considered clinicians working with child anxiety disorders.

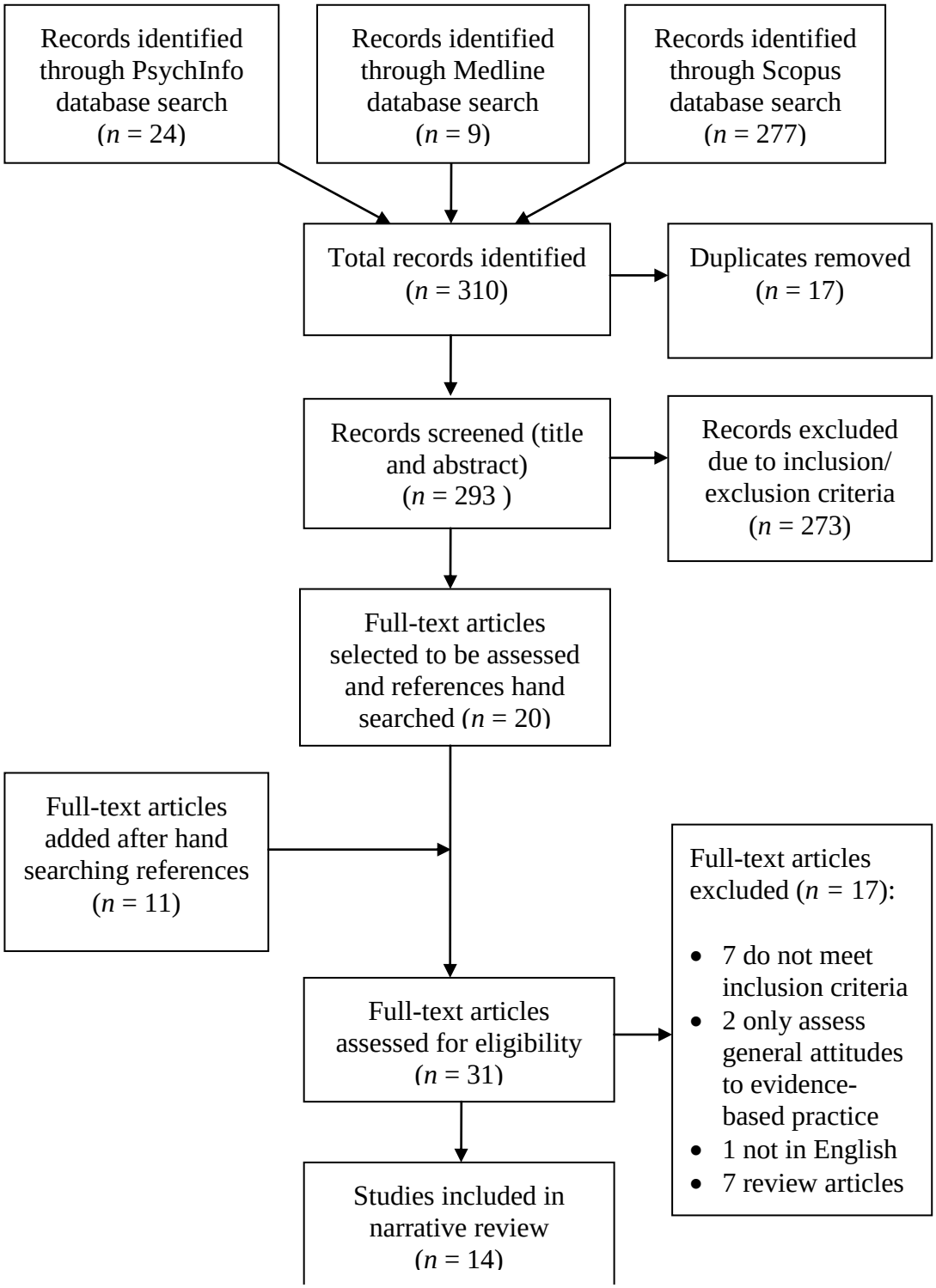

Figure 1. PRISMA diagram showing the systematic literature search procedure.

Table 1

Summary of the methodologies of included studies

Paper	Design	Participants (country)	Patient disorder	Experimental manipulation	Clinician belief or emotion assessed (outcome measure)	CBT delivery element(s) assessed (outcome measure)	Critical appraisal score (%)
Becker, Zayfert, & Anderson (2004)	Cross-sectional survey	207 psychologists and 29 members of disaster and trauma special interest group (US)	PTSD	-	Factors that influence decision to use imaginal exposure	Self-reported use of imaginal exposure	54
Beidas et al. (2014)	Quasi-experimental, pre-post, between-groups,	115 therapists working with child anxiety with minimal training in CBT for child anxiety (US)	Anxiety (child)	3 training conditions for CBT for child anxiety; training as usual, computer training, augmented (active learning) training	Beliefs about CBT for youth anxiety (CDAQ)	CBT Fidelity in performance based role play (ASCL), self-reported CBT penetration (ITAY)	50
Brown, Mountford & Waller (2014)	Cross-sectional survey	100 clinicians practicing CBT with eating disorders (UK and other countries)	Anorexia nervosa	-	Anxiety (BSI)	Percentage of clients where clinicians reported focusing on different elements of CBT	54
Deacon, Farrell et al. (2013) (study 2)	Cross-sectional survey	62 therapists working with anxiety disorders (US)	OCD	-	Beliefs about exposure (TBES)	Confident or cautious ET behaviours in response to patient vignette (ETCV)	62

Table 1 (continued)

Paper	Design	Participants (country)	Patient disorder	Experimental manipulation	Clinician belief or emotion assessed (outcome measure)	CBT delivery element(s) assessed (outcome measure)	Critical appraisal score (%)
Deacon, Lickel, Farrell, Kemp & Hipol (2013)	Cross- sectional survey	66 therapists using interoceptive exposure for panic disorder (US)	Panic disorder	-	Beliefs about the impact of cognitive reappraisal and controlled breathing on interoceptive exposure, concerns about interoceptive exposure	Self-reported delivery of interoceptive exposure and adjuncts	77
Farrell, Deacon, Kemp, Dixon & Sy (2013)	Quasi-experimental, pre-post, between-groups	53 undergraduate psychology students (US)	OCD	One group exposed to positive beliefs about ET, one group exposed to negative beliefs about ET	Beliefs about exposure (TBES). Anxiety (COWC, ASI-3, SUDS)	Self-reported exposure use (Exposure Use Questionnaire), observed exposure delivery with confederate client	60
Farrell, Kemp, Blakey, Meyer & Deacon (2016)	Quasi-experimental, pre-post, between-groups	49 mental health clinicians attending exposure workshops (not stated)	Anxiety	Two groups; standard training or enhanced, experiential training in ET for anxiety	Beliefs about exposure (TBES)	Reported planned exposure delivery in response to vignette (ETCV)	48
Harned, Dimeff, Woodcock & Contreas (2013)	Quasi-experimental, pre-post, between-groups	276 mental health clinicians, with minimal ET experience (not stated)	Anxiety	Three groups of training on ET; OLT, OLT + ME, OLT +ME + learning community	Beliefs about exposure (ATET) Anxiety (ASI)	Self-reported exposure use (Exposure Therapy Clinical Use survey), clinical proficiency in ET during role-plays	61

Table 1 (continued)

Paper	Design	Participants (country)	Patient disorder	Experimental manipulation	Clinician belief or emotion assessed (outcome measure)	CBT delivery element(s) assessed (outcome measure)	Critical appraisal score (%)
Levita, Gonzalez Salas Duhne, Girling, & Waller (2016)	Cross-sectional, correlational	32 junior therapists using CBT (UK)	Not specified	-	Anxiety - cognitive (IUS-12), physiological (SCR, HRV) and behavioural (BART)	Self-reported use of CBT techniques	69
Meyer, Farrell, Kemp, Blakey & Deacon (2014)	Cross-sectional Survey	189 mental health practitioners using ET for anxiety (US)	Anxiety	-	Beliefs about exposure (TBES) Anxiety (ASI-3)	Self-reported likelihood of excluding patients from exposure (BLES)	69
Parker & Waller (2017)	Cross-sectional Survey	204 clinicians working with anxiety disorders patients (UK)	Anxiety	-	Beliefs about CBT (NACS) Anxiety (IUS-12)	Self-reported use of CBT techniques (TMQ)	62
van Minnen, Hendricks & Olff (2010)	Quasi-experimental, between- groups	255 trauma experts (Dutch -Flemish)	PTSD	Two vignette conditions, patients' co-morbidity and preference for exposure manipulated	Beliefs about treatment credibility and perceived barriers to imaginal exposure	Clinician ratings of treatment preference for imaginal exposure and other non-exposure techniques in response to patient vignettes	61

Table 1 (continued)

Paper	Design	Participants (country)	Patient disorder	Experimental manipulation	Clinician belief or emotion assessed (outcome measure)	CBT delivery element(s) assessed (outcome measure)	Critical appraisal score (%)
Waller, Stringer & Meyer (2012)	Cross-sectional survey	100 qualified psychological therapists working with eating disordered patients (UK)	Eating disorders	-	Anxiety (BSI)	Self-reported use of specific CBT techniques	46
Whiteside, Deacon, Benito & Stewart (2016)	Cross-sectional survey	331 clinicians working with child anxiety (US)	Anxiety (child)	-	Beliefs about exposure (TBES) and beliefs about child resilience (CD-RISC:T)	Self-reported use of CBT techniques	62

Note: ASCL = Adherence and Skill Checklist (Beidas, Barmish & Kendall, 2009), ASI = Anxiety Sensitivity Index (Reiss, Peterson, Gursky & McNally, 1986), ASI-3 = Anxiety Sensitivity Index- 3 (Taylor et al., 2007), ATET = Attitudes Toward Exposure Therapy Scale (Harned, Dimeff Woodcock & Skutch, 2011), BART = Balloon Analogue Risk Task (Lejuez et al., 2002), BLES = Broken Leg Exception Scale (Meyer et al., 2014), BSI = Brief Symptom Inventory (Derogatis, 1975), CBT = Cognitive Behavioural Therapy, CDAQ = Clinician Demographics and Attitudes Questionnaire (Beidas et al., 2009), CD-RISC:T = Connor-Davidson Resilience Scale: Therapist (Connor & Davidson, 2003), COWC = Contamination Obsessions and Washing Compulsions (Subscale of the Padua Inventory-Washington State University Revision; Burns, Koertge, Formea, & Sternberger, 1996), ET = Exposure Therapy, ETCV = Exposure Therapy Case Vignette (Deacon, Farrell et al., 2013), HRV = Heart Rate Variability, ITAY = Identification and Treatment of Anxious Youth (Benjamin, Beidas, Edmunds, Cohen & Kendall, 2010), IUS-12 = Intolerance of Uncertainty- Short Form (Carleton, Norton & Asmundson, 2007), ME = Motivational Enhancement, NACS = Negative Attitudes towards CBT Scale (Parker & Waller, 2017), OCD = Obsessive compulsive disorder, OLT = Online Training, PTSD = Post-traumatic stress disorder, SCR = Skin Conductance Response, SUDS = Subjective Units of Distress Scale, TBES = Therapists Beliefs about Exposure Scale (Deacon, Farrell et al. 2013), TMQ = Therapy Methods Questionnaire (Parker & Waller, 2017), UK = United Kingdom, US = United States of America.

Quality Assessment

Details of the quality assessment can be found in Appendix B, and total quality scores are shown in Table 1. There was variation in the quality of studies, but the majority had a number of methodological flaws.

All studies clearly described their aims and main findings, and most described their main outcome measures in the methods section. However, several studies used unstandardised measures of clinician beliefs or CBT use, which might not have been valid or reliable. Almost all studies provided estimates of random variability in main outcome data, but only eight papers reported actual probability values. Samples were often poorly defined, and might not have been representative of the wider population. In the survey studies, few could provide response rates, given their sampling technique (e.g., snowball sampling), and those that did were low (<30%). No studies assessed the characteristics of participants who chose not to participate, increasing risk of self-selecting bias.

Most quasi-experimental between-group studies considered a partial list of between group confounders, excepting Beidas et al. (2014) who collapsed their groups for analysis anyway. Interventions were clearly described in all papers, but it was often unclear if the setting was representative of training staff might usually receive or if compliance with the intervention was reliable. No studies comprehensively assessed adverse outcomes. This deficit might be a particular issue for Farrell et al.'s (2013) study, as they induced negative beliefs in participants and used deception.

Only two studies of training outcomes included a follow-up (Beidas et al., 2014; Harned et al., 2013), and neither described characteristics of patients lost to follow-up. Only one study (Farrell et al., 2013) blinded study subjects to condition. Neither Beidas et al. (2014) or Farrell et al. (2016) randomised participants to condition.

All studies used appropriate statistical tests (e.g., using non-parametric statistics where data were not normally distributed), and all papers were clear where any analyses were unplanned. Most studies took some consideration of confounders in their analyses (e.g., clinician gender or profession). None of the studies reported sample size analysis, so findings might be at risk of type II errors. However, many sample sizes were large and significant results were found, which reduces this concern.

Outcomes

Key outcomes for the included studies are summarised in Table 2 and described in more detail below, in relation to the two questions posed earlier.

1. What is the relationship between clinicians' beliefs and their CBT delivery?

Most papers specifically considered the association between clinicians' beliefs about exposure and their exposure use or delivery style. Tables 1 and 2 show that clinicians' negative beliefs about exposure are related to less intention to use exposure, and that this relationship may interact with patient variables, such as patient complexity (Becker et al., 2004; Meyer et al., 2014; van Minnen et al., 2010). Conversely, clinicians' positive beliefs about the credibility of exposure were related to greater intention to use the method (van Minnen et al., 2010). However, the relationship between clinician negative beliefs and self-reported exposure use is unclear, with one study finding that negative beliefs were associated with less exposure use (Whiteside et al., 2016), and one study finding clinician beliefs did not predict their exposure use (Harned et al., 2013).

There are stronger findings about clinician negative beliefs about exposure and their exposure delivery style. There was a relationship between negative clinician beliefs and more cautious, less proficient exposure delivery (Harned et al., 2013; Deacon, Farrell et al., 2013; Deacon, Lickel et al., 2013). This cautious delivery includes clinicians selecting a less anxiety provoking exposure item, and using more

distress reduction strategies (e.g., controlled breathing), which are empirically unsupported. Furthermore findings suggest that clinicians' negative beliefs may *cause* that cautious delivery (Farrell et al., 2013; Farrell et al., 2016). Farrell et al. (2013) demonstrated causality by inducing negative and positive beliefs about exposure in participants, albeit in an undergraduate student sample. Farrell et al. (2016) found that reduction of negative beliefs mediated improvement in exposure delivery style following training. Finally, findings also suggest that positive beliefs can have a positive impact on exposure delivery, making it more confident and evidence-based (Farrell et al., 2013).

In terms of the relationship between broader beliefs about CBT and CBT delivery, it was not possible to draw any conclusions as the results are inconsistent. One study found negative beliefs about CBT were associated with less frequent use of psycho-educational and cognitive techniques (Parker & Waller, 2017), the other study found no relationship between clinician beliefs and CBT fidelity (Beidas et al., 2014).

Summary. Clinicians' negative beliefs about exposure lead to less confident exposure delivery, and might be associated with less exposure use. The reverse is indicated for clinicians' positive beliefs about exposure. However, there is not enough evidence to reach conclusions about other CBT methods.

Table 2

Summary of key findings from the included studies, indicating where clinicians' beliefs or anxiety have been found to be associated with at least one aspect of CBT delivery

Studies	Clinician beliefs				Clinician anxiety		
	Negative Beliefs about exposure	Positive beliefs about utility of unsupported adjuncts to exposure (e.g. controlled breathing)	Negative beliefs about CBT	Negative beliefs about patient resilience	Anxiety	Anxiety sensitivity	Intolerance of uncertainty
Becker et al. (2004)							
Beidas et al. (2014)							
Brown et al. (2014)							
Deacon, Farrell et al. (2013)							
Deacon, Lickel et al. (2013)							
Farrell et al. (2013)							
Farrell et al. (2016)							
Harned et al. (2014)							
Levita et al. (2016)							
Meyer et al. (2014)							
Parker & Waller (2017)							
van Minnen et al. (2010)							
Waller et al. (2012)							
Whiteside et al. (2016)							

Note: = yes, = no, = mixed findings, = not applicable.

2. What is the relationship between clinicians' emotions and their CBT delivery?

The only clinician emotion considered in the papers reviewed was anxiety. Again, some papers looked at specific exposure use and delivery, while others examined a wider range of CBT techniques (including psycho-educational, cognitive and behavioural techniques).

Findings indicate that anxiety sensitivity is related to and predicts more cautious, less proficient exposure delivery (Farrell et al., 2013; Harned et al., 2013; Meyer et al., 2014). Anxiety sensitivity also affects clinician's decisions on exposure use in hypothetical scenarios (Farrell et al., 2013; Meyer et al., 2014), but this finding might not translate into clinical use (Harned et al., 2013).

In terms of the relationship between clinician anxiety and wider CBT delivery, three studies suggest that higher clinician anxiety is related to less use of behavioural CBT techniques (Levita et al., 2016; Parker & Waller, 2017; Waller et al., 2012), while one study found no relationship (Brown et al., 2014). This difference in findings might be because Brown et al.'s study assessed CBT techniques used in the first six treatment sessions whilst the other studies assessed self-reported use of these techniques more generally. In addition Levita et al.'s paper suggests that different aspects of anxiety may be related to the use of different CBT techniques.

Summary. Clinicians' anxiety sensitivity predicts more cautious exposure delivery, and might also affect clinicians' decisions to use exposure. Different aspects of clinicians' anxiety are related to less use of behavioural techniques in CBT.

Discussion

The aim of this review was to systematically review the empirical research to explore the relationship between clinicians' beliefs and clinicians' anxiety on their CBT delivery. Findings are summarised below and their relationship to existing literature is

discussed. Recommendations are made for further research. Finally, the review considers clinical implications and limitations of the included studies and of this review.

Summary of results

The research in this area is still developing, and only fourteen papers were eligible for inclusion in this review. Furthermore, those papers had varying methodologies and outcome measures, so results should be interpreted with caution.

The findings indicate that clinicians' negative beliefs about exposure may lead to less confident exposure delivery. This cautious delivery includes clinicians selecting easier exposure items and using unsupported distress reduction techniques (e.g., controlled breathing) that might impede exposure delivery. Clinicians' negative beliefs about exposure might also be associated with less exposure use. Conversely, findings suggest clinicians' positive beliefs about exposure might lead to more confident and intense exposure delivery.

The only clinician emotion considered in the papers reviewed was anxiety, and findings are similar to those outlined above. Anxiety sensitivity predicts more cautious exposure delivery, and might also affect clinicians' decisions to use exposure. Finally, other aspects of clinicians' anxiety may be related to less use of behavioural techniques.

How do the findings relate to the wider literature base?

These findings support and expand previous literature reporting clinicians' negative beliefs about exposure and anxiety about CBT techniques (e.g., Becker et al., 2004, Deacon, Lickel et al., 2013; Olatunji et al., 2009; Turner et al., 2014; van Minnen et al. 2010), by demonstrating the relationship between these characteristics and CBT delivery. Findings that beliefs about exposure and anxiety lead to more cautious delivery also support previous work indicating that exposure is often adapted with non-evidence-based adjuncts to reduce patient distress, and provides some explanation for such clinical practice (Freiheit et al., 2004).

Perhaps most importantly, the findings of this systematic review support hypotheses about how clinician beliefs and anxiety might lead to therapist drift (Waller, 2009; Waller & Turner, 2016). The findings of this review are consistent with the suggestion that clinicians' anxiety and negative beliefs about exposure lead them to engage in safety behaviours, avoiding behavioural techniques or delivering exposure more cautiously (with distress-reducing adjuncts). This review therefore contributes to the growing literature on barriers to the dissemination of evidence-based psychological treatments, such as CBT (Shafran et al., 2009). Findings from this review suggest that one way to bridge the 'gap' between research findings and clinicians' real-life practice might be to address clinicians' own beliefs and anxiety.

Critique

There are a number of limitations of the included studies. Many were cross-sectional and correlational, which limits the generalisability of findings and prohibits causal inferences. In addition, samples were often poorly defined and at risk of self-selecting bias, so they might not be representative of the wider population of clinicians using CBT. However many studies had large samples, including a variety of mental health professionals, which might improve generalisability.

Some of the measures used to assess clinician beliefs were not validated in earlier papers, though anxiety was assessed using validated measures. Many of the measures used to assess clinicians' use of exposure or other CBT techniques were based on retrospective or prospective self-report measures, resulting in a potential for socially desirable responding or recall bias. Where exposure delivery was assessed using role plays, this approach might not represent true clinical practice. None of the studies included in this review observed actual clinical practice.

The methodology of this review can also critiqued. Strengths of this review include the systematic search of three separate databases, using comprehensive search

terms. However several studies (*n* = 11) were identified through hand-searching references, which might indicate the need for a better set of search terms and a wider set of databases. The decision to include only empirical studies in peer reviewed journals increases the risk of publication bias towards significant results. However, the decision to not search grey literature makes it more likely that the studies included were of higher quality.

Another strength of this review is the use of a quality appraisal tool to assess study quality. Critical appraisal was also second rated with good inter-rater reliability. However, it could be argued that generating total critical appraisal percentage scores was not appropriate, as this assumes that each item in the measure is weighted the same. In light of this, total percentage scores were included only to allow crude comparison between studies, and were not used to apply qualitative descriptors to studies.

Another limitation is that this review did not provide the statistical findings, such as effect sizes, reported in the included papers. This limits the ability to assess the strength of the associations reported across studies. This review also did not discuss the statistical methods used in the included studies or consider how this might affect the findings presented.

Finally, this review did not include statistical analyses of findings from the included papers. A meta-analysis was not completed due to heterogeneity in the measures and designs used across the included studies (e.g. cross-sectional, experimental, vignette-based), which would have potentially made the results of a meta-analysis uninterpretable.

Clinical Implications

The findings of this review suggest that patients being treated with CBT for eating or anxiety disorders might not receive the evidence-based behavioural elements of CBT that NICE recommends, due to the clinicians' own anxiety and negative beliefs

about exposure. Clinicians using CBT need to monitor their practice and to be sure to include behavioural elements in their delivery. They should focus on the evidence that supports delivering exposure techniques without the use of distress-reducing adjuncts, regardless of their personal beliefs or worries. Clinicians might also use these evidence-based behavioural change techniques themselves, for example by using behavioural experiments to challenge their negative beliefs about exposure. Clinicians might also benefit from exposing themselves to their anxiety about delivering exposure by using it with patients, without employing safety behaviours, and monitoring outcomes.

Regular supervision should support such evidence-based clinical practice in CBT. Supervisors should directly address clinicians' drift away from behavioural techniques and discuss these in terms of safety behaviours, supporting clinicians to explore and interrogate their underlying beliefs about exposure and to expose themselves to their anxiety.

Clinicians should also be provided with training to address their negative beliefs about exposure and anxiety. Research suggests that training is likely to be most effective when it includes: didactic information on exposure efficacy; client testimonials about its acceptability and tolerability; and experiential activities, where clinicians are exposed to their own anxiety (Deacon, Lickel et al., 2013).

Recommendations for Further Research

The main recommendation for further research is that CBT delivery should be assessed by observation of real-life clinical practice, to improve validity of the findings. Considering the relationship between clinician beliefs or emotions and other types of behavioural work (e.g., weighing in eating disorders; behavioural activation in depression) would expand the literature.

The impact of affect could be explored further by considering emotions other than anxiety (e.g., low mood). The specific impact of different facets of anxiety also

warrants further investigation. The impact of clinician beliefs about CBT on CBT delivery should also be further explored, as this remains unclear. Other clinician beliefs should also be considered, such as beliefs about patient resilience.

Finally, research should extend on the possible relationship between clinician beliefs and patient variables. For example, does patient complexity increase the salience of negative beliefs about exposure?

Conclusion

This systematic review has found that clinicians' negative beliefs about exposure and anxiety sensitivity are associated with less confident exposure delivery, and might be associated with less exposure use. When a broader range of CBT techniques are considered, aspects of clinicians' anxiety may also be related to less use of behavioural methods. These findings support hypotheses around therapist drift, which suggest that clinicians' beliefs and anxiety lead them to engage in their own safety behaviours and 'drift' away from evidence-based delivery of behavioural techniques. However, further research is required to replicate and expand findings and to address methodological flaws, particularly by observing CBT delivery in real-life clinical practice.

References

Becker, C. B., Zayfert, C., & Anderson, E. (2004). A survey of psychologists' attitudes towards and utilization of exposure therapy for PTSD. *Behaviour Research and Therapy*, *42*, 277–292. doi:10.1016/S0005-7967(03)00138-4

Becker-Haimes, E. M., Okamura, K. H., Wolk, C. B., Rubin, R., Evans, A. C., & Beidas, R. S. (2017). Predictors of clinician use of exposure therapy in community mental health settings. *Journal of Anxiety Disorders, 49,* 88–94. doi:10.1016/j.janxdis.2017.04.002

Beidas, R. S., Barmish, A. J., Kendall, P. C. (2009). Training as usual: Can therapist behavior change after reading a manual and attending a brief workshop on cognitive behavioral therapy for youth anxiety? *The Behavior Therapist, 32,* 97–101. Retrieved from: http://www.abct.org/docs/PastIssue/32n5.pdf

Beidas, R. S., Edmunds, J., Ditty, M., Watkins, J., Walsh, L., Marcus, S., & Kendall, P. (2014). Are inner context factors related to implementation outcomes in cognitive-behavioral therapy for youth anxiety? *Administration and Policy in Mental Health and Mental Health Services Research*, *41,* 788–799. doi:10.1007/s10488-013-0529-x

Benjamin, C. L., Beidas, R. S., Edmunds, J. M., Cohen, J., & Kendall, P. (2010). *Identification and treatment of anxious youth*. Unpublished instrument. Department of Psychology, Temple University, Philadelphia, Pennsylania.

Brown, A., Mountford, V., & Waller, G. (2014). Clinician and practice characteristics influencing delivery and outcomes of the early part of outpatient cognitive behavioural therapy for anorexia nervosa. *The Cognitive Behaviour Therapist*, *7*. doi:10.1017/S1754470X14000105

Burns, G. L., Koertge, S. G., Formea, G. M., & Sternberger, L. G. (1996). Revision of the Padua Inventory of obsessive-compulsive disorder symptoms. *Behaviour*

Research and Therapy, 34, 163–167. doi:10.1016/0005-7967(95)00035-6

Carleton, R. N., Norton, M. A., & Asmundson, G. J. G. (2007). Fearing the unknown: A short version of the intolerance of uncertainty scale. *Journal of Anxiety Disorders, 21,* 105-117. doi:10.1016/j.janxdis.2006.03.014

Connor, K. M., & Davidson, J. R. (2013). Development of a new resilience scale: the Connor-Davidson Resilience Scale (CD-RISC). *Depression and Anxiety, 18,* 76–82. doi:10.1002/da.10113

Deacon, B. J., Farrell, N. R., Kemp, J. J., Dixon, L. J., Sy, J. T., Zhang, A. R., & McGrath, P. B. (2013). Assessing therapist reservations about exposure therapy for anxiety disorders: The Therapist Beliefs about Exposure Scale. *Journal of Anxiety Disorders, 27,* 772–780. doi:10.1016/j.janxdis.2013.04.006

Deacon, B. J., Lickel, J. J., Farrell, N. R., Kemp, J. J., & Hipol, L. J. (2013). Therapist perceptions and delivery of interoceptive exposure for panic disorder. *Journal of Anxiety Disorders, 27,* 259–264. doi:10.1016/j.janxdis.2013.02.004

Deacon, B. J., Lickel, J. J., Possis, E. A., Abramowitz, J. S., Mahaffey, B., & Wolitzky-Taylor, K. (2012). Do cognitive reappraisal and diaphragmatic breathing augment interoceptive exposure for anxiety sensitivity? *Journal of Cognitive Psychotherapy, 26,* 257-269. doi:10.1891/0889-8391.26.3.257

Derogatis, L. R. (1975). The Brief Symptom Inventory: an introductory report. *Psychological Medicine, 13,* 595–605. doi:10.1017/S0033291700048017

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

Farrell, N. R., Deacon, B. J., Kemp, J. J., Dixon, L. J., & Sy, J. T. (2013). Do negative beliefs about exposure therapy cause its suboptimal delivery? An experimental

investigation. *Journal of Anxiety Disorders*, *27,* 763–771. doi:10.1016/j.janxdis.2013.03.007

Farrell, N. R., Kemp, J. J., Blakey, S. M., Meyer, J. M., & Deacon, B. J. (2016). Targeting clinician concerns about exposure therapy: A pilot study comparing standard vs. enhanced training. *Behaviour Research and Therapy, 85,* 53–59. doi:10.1016/j.brat.2016.08.011

Foa, E. B., & McLean, C. P. (2016). The efficacy of exposure therapy for anxiety-related disorders and its underlying mechanisms: The case of OCD and PTSD. *Annual Review of Clinical Psychology, 12,* 1-28. doi:1146/annurev-clinpsy-021815-093533

Freiheit, S. R., Vye, C., Swan, R., & Cady, M. (2004). Cognitive-behavioral therapy for anxiety: Is dissemination working? *The Behavior Therapist*, *27*, 25-32. Retrieved from: http://www.abct.org/docs/PastIssue/27n2.pdf

Harned, M. S., Dimeff, L. A., Woodcock, E. A., & Contreras, I. (2013). Predicting adoption of exposure therapy in a randomized controlled dissemination trial. *Journal of Anxiety Disorders, 27,* 754–762. doi:10.1016/j.janxdis.2013.02.006

Harned, M. S., Dimeff, L. A., Woodcock, E. A., & Skutch, J. M. (2011). Overcoming barriers to disseminating exposure therapies for anxiety disorders: A pilot randomized controlled trial of training methods. *Journal of Anxiety Disorders, 25,* 155–163. doi:10.1016/j.janxdis.2010.08.015

Hay, P. (2013). A systematic review of evidence for psychological treatments in eating disorders: 2005–2012. *International Journal of Eating Disorders, 46,* 462–469. doi:10.1002/eat.22103

Hipol, L. J., & Deacon, B. J. (2013). Dissemination of evidence-based practices for anxiety disorders in Wyoming: A survey of practicing psychotherapists. *Behaviour Modification, 37,* 170-188. doi:10.1177/0145445512458794

Hofmann, S. G., Asnaani, A., Vonk, I. J. J., Sawyer, A, T., & Fang, A. (2012). The efficacy of cognitive behavioral therapy: A review of meta-analyses. *Cognitive Therapy and Research, 36,* 427–440. doi:10.1007/s10608-012-9476-1

Koo, T. K., & Li, M. Y. (2016). A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *Journal of Chiropractic Medicine, 15,* 155-163. doi: 10.1016/j.jcm.2016.02.012

Larson, J. L., Vos, C. M., & Fernandez, D. (2013). Interventions to increase physical activity in people with COPD: systematic review. *Annual Review of Nursing Research, 31,* 297-326. doi:10.1891/0739-6686.31.297.

Levita, L., Gonzalez Salas Duhne, P. G., Girling, C., & Waller, G. (2016). Facets of clinicians' anxiety and the delivery of cognitive behavioral therapy. *Behaviour Research and Therapy, 77,* 157–161. doi:10.1016/j.brat.2015.12.015

Lejuez, C. W., Read, J. P., Kahler, C. W., Richards, J. B., Ramsey, S. E., Stuart, G. L., ... Brown, R. A. (2002). Evaluation of a behavioral measure of risk taking: The balloon analogue risk task (BART). *Journal of Experimental Psychology: Applied, 8,* 75-84. doi: 10.1037/1076-898X.8.2.75

Meyer, J. M., Farrell, N. R., Kemp, J. J., Blakey, S. M. & Deacon, B. J. (2014). Why do clinicians exclude anxious clients from exposure therapy? *Behaviour Research and Therapy, 54,* 49-53. doi: 10.1016/j.brat.2014.01.004

Moher, D., Liberati, A., Tetzlaff, J., Altman, D. G., & The PRISMA Group. (2009). Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. *PLOS Medicine, 6,* 1-6. doi:10.1371/journal.pmed1000097

National Institute for Health and Care Excellence. (2005a). *Post-traumatic stress disorder: Management, CG26*. London, UK: National Institute for Health and Care Excellence. Retrieved from: https://www.nice.org.uk/guidance/cg26

National Institute for Health and Care Excellence. (2005b). *Obsessive-compulsive disorder and body dysmorphic disorder: Treatment, CG31*. London, UK: National Institute for Health and Care Excellence. Retrieved from: https://www.nice.org.uk/guidance/cg31

National Institute for Health and Care Excellence. (2011). *Generalised anxiety disorder and panic disorder in adults: Management, CG113*. London, UK: National Institute for Health and Care Excellence. Retrieved from: https://www.nice.org.uk/guidance/cg113

National Institute for Health and Care Excellence. (2017). *Eating disorders: recognition and treatment, NG69*. London, UK: National Institute for Health and Care Excellence. Retrieved from: https://www.nice.org.uk/guidance/ng69

National Institute for Health and Care Excellence. (2018). *Depression in adults: Recognition and management, CG90*. London, UK: National Institute for Health and Care Excellence. Retrieved from: https://www.nice.org.uk/guidance/cg90

Olatunji, B. O., Cisler, J. M., & Deacon, B. J. (2010). Efficacy of cognitive behavioral therapy for anxiety disorders: a review of meta-analytic findings. *The Psychiatric Clinics of North America, 33*, 557-577. doi:10.1016/j.psc.2010.04.002

Olatunji, B. O., Deacon, B. J., & Abramowitz, J. S. (2009). The cruelest cure? Ethical issues in the implementation of exposure-based treatments. *Cognitive and Behavioral Practice, 16,* 172-180. doi:10.1016/j.cbpra.2008.07.003

Parker, Z. J., & Waller, G. (2017). Development and Validation of the Negative Attitudes towards CBT Scale. *Behavioural and Cognitive Psychotherapy, 45,* 629–646. doi:10.1017/S1352465817000170

Reiss, S., Peterson, R. A., Gursky, D. M., & McNally, R. J. (1986). Anxiety sensitivity, anxiety frequency, and the prediction of fearfulness. *Behaviour Research and Therapy. 24,* 1–8. doi:10.1016/0005-7967(86)90143-9

Schmidt, N. B., Woolaway-Bickel, K., Trakowski, J., Santiago, H., Storey, J., Koselka, M. & Cook, J. (2000). Dismantling cognitive-behavioral treatment for panic disorder: questioning the utility of breathing retraining. *Journal of Consulting and Clinical Psychology, 68,* 417-424. doi:10.1037/0022-006X.68.3.417

Shafran, R., Clark, D. M., Fairburn, C. G., Arntz, A., Barlow, D. H., Ehlers, A., … Wilson, G. T. (2009). Mind the gap: Improving the dissemination of CBT. *Behaviour Research and Therapy, 47*, 902–909. doi:10.1016/j.brat.2009.07.003

Stobie, B., Taylor, T., Quigley, A., Ewing, S., & Salkovskis, P. M. (2007). "Contents may vary": A pilot study of treatment histories of OCD patients. *Behavioural and Cognitive Psychotherapy, 35*, 273–282. doi:10.1017/S135246580700358X

Taylor, S., Zvolensky, M. J., Cox, B. J., Deacon, B., Heimberg, R. G., Ledley, D. R., ... Cardenas, S. J. (2007). Robust dimensions of anxiety sensitivity: development and initial validation of the Anxiety Sensitivity Index-3. *Psychological Assessment, 19,* 176–188. doi:10.1037/1040-3590.19.2.176

van Minnen, A., Hendriks, L., & Olff, M. (2010). When do trauma experts choose exposure therapy for PTSD patients? A controlled study of therapist and patient factors. *Behaviour Research and Therapy*, *48*, 312–320. doi:10.1016/j.brat.2009.12.003

Waller, G. (2009). Evidence-based treatment and therapist drift. *Behaviour Research and Therapy*, *47*, 119–127. doi:10.1016/j.brat.2008.10.018

Waller, G. (2016). Treatment protocols for eating disorders: Clinicians' attitudes, concerns, adherence and difficulties delivering evidence-based psychological interventions. *Current Psychiatry Reports*, *18*. doi:10.1007/s11920-016-0679-0

Waller, G., Stringer, H., & Meyer, C. (2012). What cognitive-behavioural techniques do therapists report using when delivering cognitive-behavioural therapy for the eating disorders? *Journal of Consulting and Clinical Psychology, 80,* 171-175. doi:10.1037/a0026559

Waller, G., & Turner, H. (2016). Therapist drift redux: Why well-meaning clinicians fail to deliver evidence-based therapy, and how to get back on track. *Behaviour Research and Therapy, 77,* 129-137. doi:10.1016/j.brat.2015.12.005

Whiteside, S. P. H., Deacon, B. J., Benito, K., & Stewart, E. (2016). Factors associated with practitioners' use of exposure therapy for childhood anxiety disorders. *Journal of Anxiety Disorders, 40,* 29–36. doi:10.1016/j.janxdis.2016.04.001

Appendix A

Adapted Downs and Black's Critical Appraisal Tool

Reporting

1. Is the hypothesis/aim/objective of the study clearly described?

Yes 1
No 0

2. Are the main outcomes to be measured clearly described in the Introduction or Methods section?

If the main outcomes are first mentioned in the Results section, the question should be answered no.

Yes 1
No 0

3. Are the characteristics of the participants included in the study clearly described?

In cohort studies and trials, inclusion and/or exclusion criteria should be given.

Yes 1
No 0

4.Are the training interventions clearly described?

Yes 1
No 0

5. Are the distributions of principal confounders in each group of participants to be compared clearly described?

A list of principal confounders is provided.

Yes 2
Partially 1
No 0

6. Are the main findings of the study clearly described?

Simple outcome data should be reported for all major findings so that the reader can check the major analyses and conclusions .(This question does not cover statistical tests which are considered below).

Yes 1

No 0

7. Does the study provide estimates of the random variability in the data for the main outcomes?

In non-normally distributed data the inter-quartile range of results should be reported. In normally distributed data the standard error, standard deviation or confidence intervals should be reported around the therapist effect. If the distribution of the data is not described, it must be assumed that the estimates used were appropriate and the question should be answered yes.

Yes 1

No 0

8. Have all important adverse events that may be a consequence of the training intervention been reported?

This should be answered yes if the study demonstrates that there was a comprehensive attempt to measure adverse events. (A list of possible adverse events is provided).

Yes 1

No 0

9. Have the characteristics of participants lost to follow-up been described?

This should be answered yes where there were no losses to follow-up or where losses to follow-up were so small that findings would be unaffected by their inclusion. This should be answered no where a study does not report the number of patients lost to follow-up.

Yes 1

No 0

10. Have actual probability values been reported (e.g. 0.035 rather than <0.05) for the main outcomes except where the probability value is less than 0.001?

Yes 1

No 0

External validity

All the following criteria attempt to address the representativeness of the findings of the study and whether they may be generalised to the population from which the study subjects were derived.

11. Were the subjects asked to participate in the study representative of the entire population from which they were recruited?

The study must identify the source population for participants and describe how the participants were selected. Participants would be representative if they comprised the entire source population, or a random sample. Random sampling is only feasible where a list of all members of the relevant population exists. Where a study does not report the proportion of the source population from which the participants are derived, the question should be answered as unable to determine.

Yes	1
No	0
Unable to determine	0

12. Were those subjects who were prepared to participate representative of the entire population from which they were recruited?

The proportion of those asked who agreed should be stated. Validation that the sample was representative would include demonstrating that the distribution of the main confounding factors was the same in the study sample and the source population.

Yes	1
No	0
Unable to determine	0

13. Were the staff, places, and facilities where the participants received the training intervention, representative of training the majority of clinicians receive?

For the question to be answered yes the study should demonstrate that the training was representative of that in use in the source population.

Yes	1
No	0
Unable to determine	0

Internal validity –bias

14. Was an attempt made to blind participants to the training intervention they have received?

Yes	1
No	0
Unable to determine	0

15. Was an attempt made to blind those measuring the main outcomes of the training?

Yes	1
No	0
Unable to determine	0

16. If any of the results of the study were based on "data dredging", was this made clear?

Any analyses that had not been planned at the outset of the study should be clearly indicated. If no retrospective unplanned subgroup analyses were reported, then answer yes.

Yes	1
No	0
Unable to determine	0

17. In trials and cohort studies, do the analyses adjust for different lengths of follow-up of patients?

Where follow-up was the same for all study participants the answer should be yes. If different lengths of follow-up were adjusted for by, for example, survival analysis the answer should be yes. Studies where differences in follow-up are ignored should be answered no.

Yes	1
No	0
Unable to determine	0

18. Were the statistical tests used to assess the main outcomes appropriate?

The statistical techniques used must be appropriate to the data. For example non-parametric methods should be used for small sample sizes. Where little statistical analysis has been undertaken but where there is no evidence of bias, the question should be answered yes. If the distribution of the data (normal or not) is not described it must be assumed that the estimates used were appropriate and the question should be answered yes.

Yes	1
No	0
Unable to determine	0

19. Was compliance with the training intervention assessed?

Where there was non-compliance with the allocated treatment or where there was contamination of one group, the question should be answered no. For studies where the effect of any misclassification was likely to bias any association to the null, the question should be answered yes.

Yes	1
No	0
Unable to determine	0

20. Were the main outcome measures used accurate (valid and reliable)?

For studies where the outcome measures are clearly described, the question should be answered yes. For studies which refer to other work or that demonstrates the outcome measures are accurate, the question should be answered as yes.

Yes	1
No	0
Unable to determine	0

Internal validity -confounding (selection bias)

21. Were the participants in different training intervention groups recruited from the same population?

The question should be answered unable to determine for cohort studies where there is no information concerning the source of patients included in the study.

Yes	1
No	0
Unable to determine	0

22. Were study subjects in different training groups recruited over the same period of time?

For a study which does not specify the time period over which patients were recruited, the question should be answered as unable to determine.

Yes	1
No	0
Unable to determine	0

23. Were participants randomised to training groups?

Studies which state that participants were randomised should be answered yes except where method of randomisation would not ensure random allocation. For example alternate allocation would score no because it is predictable.

Yes 1

No 0

Unable to determine 0

24. Was the randomised training condition assignment concealed from both participants and trainers until recruitment was complete and irrevocable?

All non-randomised studies should be answered no. If assignment was concealed from patients but not from staff, it should be answered no.

Yes 1

No 0

Unable to determine 0

25. Was there adequate adjustment for confounding in the analyses from which the main findings were drawn?

This question should be answered no for randomised studies if: the main conclusions of the study were based on analyses of treatment rather than intention to treat; the distribution of known confounders in the different treatment groups was not described; or the distribution of known confounders differed between the treatment groups but was not taken into account in the analyses. In non-randomised studies if the effect of the main confounders was not investigated or confounding was demonstrated but no adjustment was made in the final analyses the question should be answered as no.

Yes 1

No 0

Unable to determine 0

26. Were losses of participants to follow-up taken into account?

If the numbers of patients lost to follow-up are not reported, the question should be answered as unable to determine. If the proportion lost to follow-up was too small to affect the main findings, the question should be answered yes.

Yes 1

No 0

Unable to determine 0

Power

27. Did the study have sufficient power to detect a clinically important effect where the probability value for difference being due to chance is less than 5%?

The question should be scored yes if the author reports statistical power and power was attained. Where a sample size analysis was not conducted the question should be scored unable to determine.

Yes	1
No	0
Unable to determine	0

Appendix B

Table Showing Critical Appraisal for Included Studies

Items:	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	Total score (%)
Becker et al. (2004)	1	0	0			1	1			0	1	0				1		1		0					1		0	54
Beidas et al. (2014)	1	0	1	1	0	1	1	0	0	1	0	0	0	0	1	1	1	1	0	0	1	0	0	1	1	1	0	50
Brown et al. (2014)	1	1	0			1	1			0	0	0				1		1		0					1		0	54
Deacon, Farrell et al. (2013)	1	1	0			1	1			1	0	0				1		1		1					0		0	62
Deacon, Lickel et al. (2013)	1	1	1			1	1			1	1	0				1		1		0					1		0	77
Farrell et al. (2013)	1	1	1	1	1	1	1	0		1	0	0	0	1	0	1		1	1	1	1	0	1	0	0		0	60
Farrell et al. (2016)	1	1	0	1	1	1	1	0		1	0	0	0	0	0	1		1	1	1	0	0	0	0	1		0	48
Harned et al. (2014)	1	1	1	1	1	1	1	0	0	1	0	0	1	0	0	1	1	1	0	1	1	1	1	0	1	0	0	61
Levita et al. (2016)	1	1	1			1	1			0	0	0				1		1		1					1		0	69
Meyer et al. (2014)	1	1	1			1	1			0	1	0				1		1		1					0		0	69
Parker & Waller (2017)	1	1	0			1	1			1	0	0				1		1		0					1		0	62
Van Minnen et al. (2010)	1	1	0	1	1	1	1	0		1	1	0		0	0	1		1		0	1	1	1	0	1		0	61
Waller et al. (2012)	1	1	0			1	0			0	0	0				1		1		0					1		0	46
Whiteside et al. (2016)	1	1	0			1	1			0	1	0				1		1		0					1		0	62

This page is left intentionally blank

Part Two: Research Report

Clinician weighing behaviours with eating-disordered patients: the role of patient and clinician characteristics

Abstract

Objective: This experimental study aimed to explore the impact of patient and clinician characteristics on clinicians' intended weighing behaviours when using CBT for eating disorders.

Method: Seventy four clinicians using CBT with eating-disordered patients completed an online survey. Clinicians rated their likelihood of engaging in six weighing behaviours in response to patient vignettes where the patient's diagnosis and level of distress were manipulated. Clinicians completed questions regarding demographic information and beliefs about open weighing. Clinicians' anxiety, firm empathy and likelihood of making 'broken leg exceptions', were assessed using validated psychometric measures.

Results: Clinicians' intended weighing behaviours differed according to patient characteristics (diagnosis and distress). Clinicians' negative beliefs about open-weighing were associated with less compliant weighing behaviours and this appeared to be moderated by patient characteristics (diagnosis and weight-related distress). Conversely, positive beliefs about open-weighing were associated with more compliant weighing behaviours. Clinicians' tendency to make broken leg exceptions was associated with less compliant weighing behaviours, regardless of patient characteristics.

Conclusions: In CBT for eating disorders, clinicians' intended weighing behaviours might be influenced by their negative beliefs about open weighing and tendency to make broken leg exceptions, particularly in the face of patient characteristics such as diagnosis and weight-related distress. The potential clinical implications for patients who do not receive evidence-based interventions, and recommendations for clinician training and supervision around open weighing are highlighted. Strengths and limitations of the study and avenues for further research are discussed, such as observing clinicians' weighing behaviours in clinical practice.

Practitioner Points

1. Clinician's beliefs can influence their decisions about weighing patients when using CBT for eating disorders, and these beliefs may interact with patient characteristics.
2. Clinicians should be prepared to openly weigh all patients in each session of CBT for eating disorders, and acknowledge that a decision to do otherwise reflects their personal beliefs and not the evidence-base.
3. Clinicians should be prepared to challenge their negative beliefs about open weighing by putting it into practice and monitoring outcomes.
4. More research is needed to assess whether the findings of this vignette-based, self-report survey translate into clinician's actual weighing behaviours in real-life clinical practice.

Introduction

Eating Disorders

Eating disorders are serious mental health conditions, where sufferers experience negative beliefs about themselves and their eating, body shape and weight (National Institute for Health and Care Excellence, NICE, 2017). According to the Diagnostic and Statistical Manual, DSM-5 (American Psychiatric Association, APA, 2013), the main categories of eating disorders are anorexia nervosa, bulimia nervosa, binge eating disorder and other specified feeding or eating disorder. These eating disorders share common characteristics, including disturbed body image and fear of weight gain. Anorexia nervosa is further characterised by calorie restriction leading to significantly low body weight. Bulimia nervosa is characterised by recurrent episodes of binge eating, followed by compensatory behaviours (e.g., vomiting). Binge eating disorder is also characterised by recurrent binge eating but without compensatory behaviours. Finally, individuals diagnosed with other specified feeding or eating disorder present with some symptoms of an eating disorder but do not meet full diagnostic criteria.

In 2015, Beat estimated that the UK prevalence of eating disorders was between 600,000-725,000, and increasing yearly. Eating disorders have significant personal and societal, health and economic costs, and can be life-long conditions if left untreated (Beat, 2015; NICE, 2017). Fortunately however, eating disorders often respond to treatment, with over 45% of individuals with anorexia nervosa or bulimia nervosa making a full recovery (Steinhausen, 2002; Steinhausen, & Weber, 2009). Therefore, it is vital that evidence-based, effective treatment is offered and delivered to eating disorder patients.

Treatments

The NICE (2017) guidelines for eating disorders recommend that adults with anorexia nervosa should be treated using either: individual eating-disorder-focused

cognitive behavioural therapy (CBT-ED); Maudsley anorexia nervosa treatment for adults (MANTRA); or specialist supportive clinical management (SSCM). For bulimia nervosa in adults, CBT-ED is the only recommended treatment. CBT-ED currently has the strongest evidence base for the psychological treatment of adults with eating disorders (Fairburn & Harrison, 2003; Hay, 2013).

The Role of Weighing

The recommended therapies described above share some treatment elements, including the firm recommendation that regularly weighing the patient is a key treatment component. In addition, the manuals for CBT-ED, MANTRA, and SSCM instruct clinicians to weigh their patient openly, such that the patient is informed about their weight (Waller & Mountford, 2015).

Waller and Mountford (2015) suggest four reasons for regular, open weighing in CBT-ED. First, it allows the clinician to ensure patient safety (e.g., when severe weight loss becomes life-threatening). Second, weight can be a more valid indication of changes in eating patterns than self-report diary measures. Third, open weighing acts as a behavioural intervention, reducing the patient's anxieties around weight gain (i.e., exposure). Finally, open weighing can be used as part of a behavioural experiment that addresses the patient's 'broken cognition' - the faulty belief that eating a small number of calories will result in an unrealistically large weight gain.

Conversely, blind weighing means that the patient is weighed but is not told their weight. While this practice is not recommended in any evidence-based therapy, blind weighing is widely implemented in practice. Forbush, Richardson and Bohrer (2015) surveyed clinicians about their preference to weigh openly or blindly, and found that only 47% endorsed open weighing. Indeed, many clinicians do not weigh their patients at all (e.g., Cowdrey & Waller, 2015; Mulkens, de Vos, de Graff & Waller, under consideration; Waller Stringer and Meyer, 2012). For example, in Waller et al.'s

(2012) study of adherance to specific CBT techniques, fewer than 40% of CBT clinicians reported weighing eating disorder patients routinely, and 17% were not weighing their patients at all. This deficit is clearly concerning, given the benefits of open weighing and the risks of failing to weigh.

Why are eating disorder clinicians not weighing their patients appropriately or at all?

One reason that might explain why some clinicians are not following recommended practice is that they are unaware of the protocols (Addis & Krasnow, 2000). However, Waller et al. (2013) found that 92% of clinicians treating eating disorders were aware of manuals. However, only half of these clinicians used manuals 'often', suggesting there are other reasons for failing to follow evidence-based protocols. Other studies have found similarly low adherence to evidence-based treatment manuals in eating disorders (e.g., Tobin, Banker, Weisberg & Bowers, 2007). These findings are concerning, as clinicians who use manuals are more likely to have patients who achieve NICE guidelines on weight gain (Brown, Mountford & Waller, 2014).

The failure to deliver evidence-based treatment optimally, despite having the resources to do so, has been termed 'therapist drift' (Waller, 2009). Therapist drift in CBT has been found to particularly apply to more behavioural, exposure-based techniques (such as open weighing), and these techniques are frequently avoided by clinicians (Waller & Turner 2016). Therapist drift has been linked to clinicians' personalities, knowledge, beliefs, attitudes, emotions (particularly anxiety), interpersonal context and safety behaviours (Waller & Turner 2016). Several of these clinician characteristics might therefore influence clinicians' weighing behaviours when working with eating disorders.

Beliefs. The beliefs that clinicians hold about open weighing may be an important factor. Clinicians might believe that open weighing damages the therapeutic alliance, that it is not an effective component of CBT-ED, that they can judge patient weight by eye, or that someone else should weigh the patient (Waller & Mountford, 2015). The clinician may also make 'broken leg exceptions' (Meehl, 1973), citing a range of 'reasons' why open weighing is not applicable to this particular client (e.g., the patient's distress or co-morbidity), despite a lack of evidence to justify such exceptions (Meyer, Farrell, Kemp, Blakey & Deacon, 2014).

Emotions. Clinicians' anxiety may also play a role. Waller et al. (2012) found that clinicians with higher levels of anxiety reported using fewer exposure-based techniques with eating disordered patients. Turner, Tatham, Lant, Mountford and Waller's (2014) research suggests that this anxiety might be underpinned by a higher intolerance of uncertainty. This theory is similar to Meehl's (1973) suggestion that some clinicians have a 'spun-glass theory of the mind' of their patient, meaning that clinicians fear that they might 'break' patients who are too fragile to cope if asked to implement changes. Thus, an anxious therapist is more likely to engage in their own safety behaviours, not requiring the patient to be weighed. While this reduces the clinician's anxiety in the short-term, it is likely to make the clinician less effective in the long-term.

An alternative to this fear of patient fragility is a concern by clinicians that patients will resist key treatment elements. Waller et al. (2012) suggested that anxious clinicians are more likely to fear resistance from the client, so they back off from using behavioural techniques. In response to such fears, Wilson, Fairburn and Agras (1997) discussed the need for 'firm empathy' when delivering CBT-ED. This stance requires the clinician to have both an empathic understanding of the patient and firm boundaries around what needs to be done. A lack of firm empathy might be related to failure to follow weighing guidelines.

Clinician and patient demographics. Clinicians' characteristics, such as age, experience and professional discipline, may also influence their weighing practices. For example, Forbush et al. (2015) found that more experienced clinicians were more likely to weigh openly. However, Brown et al. (2014) found that older clinicians were more likely to continue seeing patients who refused to be weighed, regardless of years of experience.

The clinician characteristics outlined are likely to interact with patient characteristics. For example, Forbush et al. (2015) found that clinicians reported being less likely to tell patients their weight when the patient appeared less emotionally stable. Diagnosis also influenced weighing practices - clinicians were more likely to use blind weighing when patients were diagnosed with anorexia nervosa rather than with any other eating disorder.

Clinical Implications

The implications of clinicians failing to weigh their eating-disordered patients include reduced safety and reduced understanding of their patient's eating behaviours. Crucially, failing to weigh openly reduces opportunities to address the ‘broken cognition’ in therapy and to reduce the patient’s anxiety around weight gain via exposure. These are key cognitive and behavioural components of evidence-based CBT for eating disorders, and omitting them is likely to reduce treatment effectiveness.

The current study is designed to further our understanding of why clinicians fail to use exposure techniques in CBT, through using a robust experimental design. The results of this study could be used to inform staff training and supervision directed at promoting adherence to evidence-based CBT-ED, giving patients the best chance of recovery. Farrell, Deacon, Dixon and Lickel (2013) and Waller and Turner (2016) present some preliminary ideas on how training can be used to encourage staff adherence to evidence-based CBT. This research project could contribute to the

development of these training proposals, by clarifying specific variables to target to improve clinicians' weighing behaviour compliance.

Summary

Open weighing is a vital component of evidence-based CBT-ED. However, such weighing is not routinely practiced, potentially due to both patient and clinician characteristics. The research in this area is still emerging. To date, it has been correlational in nature, and based on clinicians' retrospective self-reports about their weighing practices. Such reports are clearly potentially biased, and more robust methodologies are needed.

This study therefore aims to expand the literature by investigating the impact of both patient and clinician characteristics on clinicians' weighing behaviours, using an experimental design and focusing on variables drawn from the existing literature. The experimental design will be based on patient vignettes. It will not focus on how often clinicians report intentions to weigh, but on how those intended weighing behaviours change when the clinical condition is manipulated via the vignettes presented.

Aims

The aims of this study are:

1. To investigate whether patient characteristics impact on clinicians' planned weighing behaviours (defined as likelihood of: weighing, open weighing, relying on patient self-reported weight, judging patient weight by eye, delaying weighing, or asking someone else to weigh the patient).
2. To investigate whether clinicians' characteristics impact on their planned weighing behaviours.

Hypotheses

1. Clinicians' intended weighing behaviours will differ according to patient diagnosis (anorexia nervosa or bulimia nervosa).
2. Clinicians' intended weighing behaviours will be more compliant with guidelines when working with patients with low levels of distress than with patients with high levels of distress.
3. Clinicians' intended weighing behaviours will be less compliant with guidelines when the patient is experiencing high levels of distress related directly to weight than when they are experiencing high levels of general distress.
4. Clinicians' characteristics (age, gender, experience, discipline, anxiety, firm empathy, beliefs about open weighing, and 'broken leg exceptions') will be associated with their intended weighing behaviours.

Method

Ethics

Ethical approval for this project was granted by the University of Sheffield's Department of Psychology Research Ethics Committee (Appendix A).

Design

This exploratory study employed a quantitative, within-subjects, experimental design. In an online survey, self-report questionnaires and a series of vignettes varying clinical content were used to assess the impact of patient and clinician characteristics on clinicians' anticipated weighing behaviours.

Participants

Power analysis. A priori sample size analysis was conducted using G*Power (Faul, Erdfelder, Lang, & Buchner, 2007) to determine the sample size required to prevent type II errors (Appendix B). The primary aim of the study was that patient diagnosis and distress level would interact to influence clinician behaviour, and this analysis was used as the basis for the sample size analysis. Assuming a two-way,

within-subject ANOVA, an effect size of $f = 0.25$, and a significance level of $p = .05$, a total sample size of 28 would be required to achieve 80% power. If the effect size were lower ($f = .20$), then the sample size needed would be 42. A total of 74 participants were recruited, ensuring that all statistical tests were adequately powered.

Recruitment. Participants were recruited based on an opportunistic snow-ball sampling method. Clinicians using CBT for eating disorders with adults were contacted using the supervisor's extensive contacts in national and international eating disorder services. Clinicians were sent an email containing a hyperlink to the online survey inviting them to participate and to forward the email to their colleagues (Appendix C).

Characteristics. The sample consisted of 74 clinicians using CBT for eating disorders with adults. Eight clinicians were male (10.8%), 65 were female (87.8%), and one participant chose not to disclose their gender. Clinicians ranged in age from 24 to 64 years, with a mean age of 38.65 years ($SD = 10.27$). Their years of experience using CBT for eating disorders ranged from 0 to 27 years, with a mean experience of 7.30 years ($SD = 6.71$).

Clinicians described themselves as psychologists ($n = 36$, 48.6%), CBT therapists ($n = 12$, 16.2%), psychotherapists ($n = 6$, 8.1%), nurses ($n = 5$, 6.8%), or 'other' professionals ($n = 15$, 20.3% - including assistant psychologists, psychiatrists, social workers, an occupational therapist and a dietitian). Fifty-eight clinicians (78.4%) were qualified in their core profession, with 16 (21.6%) in training. The sample was international, with clinicians practicing in the following countries; UK ($n = 35$, 47.3%), Canada ($n = 18$, 24.3%), USA ($n = 7$, 9.5%), Italy ($n = 7$, 9.5%), New Zealand ($n = 2$, 2.7%), Australia ($n = 1$, 1.4%), and 'other' ($n = 4$, 5.4% - including Norway, Sweden and the Isle of Man).

Measures

An online survey was developed using Qualtrics software.

Independent variables.

***Clinician demographics*.** The first items of the survey asked participants to indicate their age, gender, profession, country of practice, years of experience using CBT for eating disorders, and whether they were qualified or in training in their core profession.

Clinician cognitive and emotional characteristics. Clinicians' anxiety, firm empathy, beliefs about open weighing, and broken leg exceptions were assessed.

Anxiety. Intolerance of uncertainty has been proposed as a trans-diagnostic feature underlying anxiety disorders (Carleton et al., 2012). Clinician anxiety was therefore assessed using the Intolerance of Uncertainty Scale-Short Form (IUS-12; Carleton, Norton, & Asmundson, 2007; Appendix D). This 12-item self-report measure assesses two factors: prospective and inhibitory anxiety. Prospective anxiety is described as the anticipation of uncertainty, and inhibitory anxiety is described as inaction when faced with uncertainty (McEvoy & Mahoney, 2011). Reponses are based on a five-point Likert scale. Carleton et al. (2007) report good correlations with the original measure and related measures of anxiety, and excellent internal consistency overall (α = .91) and for both factors (α = .85).

In this study the IUS-12 showed good internal consistency for the prospective anxiety (α = .84) and inhibitory anxiety subscales (α = .88). Clinicians' scores were slightly below the mean for a non-clinical sample (Carleton et al., 2007) for prospective anxiety (M = 14, SD = 4.56) and inhibitory anxiety (M = 7.55, SD = 3.19).

Beliefs about open weighing were assessed by asking clinicians to rate their agreement with six statements, using a 7-point Likert scale (Appendix E). Four statements presented negative beliefs about open weighing, and two presented positive beliefs about open weighing. These questions were developed by the authors based on

clinical experience, so no previous psychometrics were available. They were not used as a scale in this study, so no total score was generated.

In this study, these questions demonstrated adequate internal consistency for positive beliefs (α = .76) and negative beliefs (α = .71), though it is recognised that the low number of items makes Cronbach's alpha unreliable. Clinicians' mean scores were lower for negative beliefs, ranging from 1.50 (*SD* = .94) to 2.72 (*SD* = .1.26), and higher for positive beliefs, ranging from 5.74 (*SD* = .86) to 6.45 (*SD* = 1.01).

'Broken leg exceptions'. The Broken Leg Exception Scale (BLES, Meyer et al., 2014) assesses various patient characteristics that might lead clinicians to exclude a patient from exposure tasks. The authors adapted this measure to make it specific to open weighing as the exposure, and to remove items irrelevant to this study (e.g., 'the client's feared situation(s) are difficult to recreate in real life'). The adapted BLES was termed The Broken Leg Exception Scale for Open Weighing (BLES-OW), and was a 16-item self-report scale (Appendix F). Items were scored on a 4-point Likert scale. Scores could range from 0 to 48 and a higher score represented more broken leg exceptions. In this study the BLES-OW demonstrated excellent internal consistency (α = .95). Clinicians' mean score on the BLES-OW was 5.46 (*SD* = 7.47).

Firm empathy was assessed using the Firm Empathy Questionnaire (FEQ; McAdam Freud & Waller, under consideration; Appendix G). This 16 item self-report questionnaire assessed clinicians' levels of empathy and firmness. Responses were scored on a 7-point Likert scale. The FEQ has been recently developed by the research supervisor, and full psychometric details are not yet available. Factor analysis has shown two internally consistent scales of firmness and empathy among a non-patient sample of males and females. In this study the firmness scale showed poor internal consistency (α = .63) and the empathy factor demonstrated adequate internal

consistency (α = .71). Clinicians' mean scores were 5.49 (*SD* = .53) for firmness and 5.22 (*SD* = .51) for empathy.

Experimental variable. A core clinical vignette was developed by the authors to provide basic information about an adult female patient presenting to eating disorder services for CBT (Appendix H). Six vignettes were created by varying the clinical material, with three clients described as being diagnosed with anorexia nervosa and three with bulimia nervosa. For each diagnosis, one patient was described as calm, one was described as distressed about general life events, and one was described as distressed about their weight. These vignettes were reviewed and approved by two consultant clinical psychologists working in eating disorder services.

Dependent variables. Following the presentation of each vignette, clinicians were asked to rate the likelihood that they would engage in each of six different weighing behaviours with the patient (Appendix I). The weighing behaviours were selected by the authors based on clinical experience, and represented two guideline compliant behaviours (weigh the patient; weigh the patient and tell them their weight) and four behaviours that were non-compliant with guidelines (rely on patient self-reported weight; ask someone else to weigh; judge patient weight by eye; delay weighing until another session; Waller & Mountford, 2015). Participants rated their anticipated likelihood of engaging in each weighing behaviour on an 11-point Likert-scale, using 10% intervals from 0% to 100% likelihood.

Procedure

Participants were directed to the online survey through a hyperlink in the recruitment email. To ensure informed consent, the first survey page was an information sheet, detailing the purpose of the study, what it involved, data confidentiality and the process to raise any concerns (Appendix J). Participant data were fully anonymised, and participants were asked to generate a unique code to allow them to withdraw their data.

No participants requested to withdraw their data during this study. To proceed with the survey, participants were asked to indicate that they understood their right to withdraw and that the data would be confidential, and that they consented to completing the survey. Participants who did not consent were directed to the end of the survey.

Consenting participants were able to access the survey and were asked to complete questionnaires about their demographic information, as described above. Participants were shown six patient vignettes, varying patient diagnosis and level of distress, presented in a randomised order to exclude order effects. Participants were asked to read each vignette and then answer questions about their anticipated weighing behaviours for that patient. Following the vignettes, participants were asked to complete the following self-report questionnaire measures assessing their cognitive and emotional characteristics: questions about beliefs about open weighing; IUS-12; FEQ; and BLES-OW.

Participants were required to answer each question in order to progress to the next question. On completing the survey, participants were shown a debriefing page thanking them for taking part and explaining the purpose of the study (Appendix K). This page included a reminder of the researcher's contact details in case of queries or concerns. It also briefly outlined the importance of weighing patients when using CBT for eating disorders and advised clinicians to seek supervision regarding this if they needed further support.

Data Analysis

Data were downloaded from Qualtrics into an Excel spreadsheet, which was subsequently transferred to IBM SPSS Version 23 for data analysis. Histograms were generated to assess the distribution of data. A substantial amount of the data were found to violate the assumption of normality, particularly for the weighing behaviors. In light of this, non-parametric statistics were utilized.

A minimum significance level of $p < .05$ was used in the interpretation of Wilcoxon signed-rank analyses. For most other analyses, where there were considerably more variables, the significance level was adjusted to $p < .01$ to reduce the likelihood of incorrectly rejecting the null hypothesis. Bonferroni corrections were not used, to reduce the risk of type II error.

Descriptive analyses. Participant demographics were analysed using descriptive statistics. Means and standard deviations were calculated for continuous data, and frequencies and percentages were calculated for categorical data. Means and standard deviations (rather than medians) are also reported throughout for weighing behaviours, to aid interpretation of the findings. Weighing behaviours were analysed according to their Likert scale scoring of 1-11, then means were converted into percentage likelihood of using the behaviour. For example, a mean score of 10.2928 was reported as a mean 92.93% likelihood of engaging in the weighing behaviour.

Inferential analysis. To test hypothesis 1, Clinicians' intended weighing behaviour scores were averaged across levels of patient distress to provide mean scores for six weighing behaviours for patients with anorexia nervosa or bulimia nervosa. A series of Wilcoxon signed-rank tests were then used to compare clinicians' intended likelihood of engaging in the six weighing behaviours when patients were diagnosed with anorexia nervosa or bulimia nervosa. Effect sizes were calculated for significant findings using Tau (Rosenthal, 1991).

To test hypotheses 2 and 3, a series of Friedman's ANOVAs were conducted to compare clinicians' intended likelihood of engaging in six weighing behaviours across three levels of patient distress (calm, general distress, or weight-related distress). For significant results, follow-up Wilcoxon signed-rank tests were used to compare groups and effect sizes were calculated.

To test hypothesis 4, a range of tests were used to assess the impact of clinician demographic characteristics. To compare clinicians' intended weighing behaviours by clinician gender, Mann-Whitney tests were conducted separately for anorexia nervosa and bulimia nervosa patient vignettes. To compare weighing behaviour across professions, professions were grouped into three categories to ensure a high enough *n* to run the analysis. Kruskal-Wallis tests were then conducted separately for anorexia nervosa and bulimia nervosa patient vignettes, comparing clinician's intended weighing behaviours between three professional groupings (psychologist, CBT therapist or other profession). To assess the relationships between clinician's intended weighing behaviours and their age and years of experience using CBT for eating disorders, Spearman's correlation coefficients were conducted separately for anorexia nervosa and bulimia nervosa patients' vignettes.

To assess the relationship between clinicians' cognitions and emotions and their anticipated weighing behaviours, a series of Spearman's correlation coefficients were conducted separately for anorexia nervosa and bulimia nervosa patient vignettes. These correlations assessed the relationships between clinician's intended weighing behaviours and their: beliefs about open weighing; prospective and inhibitory anxiety IUS subscale scores; BLES-OW total scores; and firmness and empathy FEQ subscale scores.

Results

Overall, clinicians reported weighing behaviour intentions that were in the direction of guideline compliance, with a higher intended likelihood of weighing (*M* = 91.62%, *SD* = 1.55) or openly weighing the patient (*M* = 92.28% , *SD* = 1.32) and lower likelihood of relying on self-reported weight (*M* = 7.08%, *SD* = 1.77), judging weight by eye (*M* = 7.08%, *SD* = 1.77), delaying weighing (*M* = 11.03%, *SD* =1.54), or asking someone else to weigh (*M* = 5.47%, *SD* =1.79).

Hypothesis 1: Clinicians' intended weighing behaviours will differ according to patient diagnosis (anorexia nervosa or bulimia nervosa).

To test hypothesis 1, a series of Wilcoxon signed-rank tests compared clinicians' intended likelihood of engaging in six weighing behaviours when patients were diagnosed with anorexia nervosa or bulimia nervosa (Table 1).

Clinicians were significantly more likely to anticipate weighing the patient if they were diagnosed with anorexia nervosa than if they were diagnosed with bulimia nervosa. Conversely, clinicians were significantly more likely to anticipate relying on the patient's self-reported weight if they were diagnosed with bulimia nervosa. Clinicians' intended likelihood of openly weighing the patient, judging patient weight by eye, delaying weighing or asking someone else to weigh the patient, did not significantly differ according to patient diagnosis.

These findings support hypothesis 1, as some of the clinician's intended weighing behaviours did differ according to patient diagnosis.

Table 1

Comparing clinicians' intended likelihood of engaging in different weighing behaviours for patients diagnosed with anorexia nervosa or bulimia nervosa

Weighing Behaviour	Patient Diagnosis				Wilcoxon signed-rank		
	Anorexia nervosa		Bulimia nervosa		*z*	*p* (2-tailed)	*r*
	M	*SD*	*M*	*SD*			
Weigh patient	92.93	(1.42)	90.32	(1.83)	-2.14	.03	-.18
Open weigh	92.02	(1.43)	92.64	(1.31)	0.76	*ns*	-
Rely on self-reported weight	5.99	(1.65)	8.31	(1.97)	2.67	.01	.22
Judge weight by eye	11.62	(2.28)	10.99	(2.25)	-0.70	*ns*	-
Delay weighing	10.50	(1.84)	11.55	(1.96)	1.22	*ns*	-
Ask someone else to weigh	5.95	(1.94)	5.00	(1.67)	-1.29	*ns*	-

Note. ns = non-significant. Significance level set at $p < .05$.

Hypothesis 2: Clinicians' intended weighing behaviours will be more compliant with guidelines when working with patients with low levels of distress than with patients with high levels of distress.

Hypothesis 3: Clinicians' intended weighing behaviours will be less compliant with guidelines when the patient is experiencing high levels of distress related directly to weight than when they are experiencing high levels of general distress.

To test hypotheses 2 and 3, a series of Friedman's ANOVAs compared clinicians' intended likelihood of engaging in six weighing behaviours across three levels of patient distress (Table 2). For significant results, follow-up Wilcoxon tests were used to compare groups.

Clinicians were significantly more likely to anticipate weighing or openly weighing patients who were calm compared to patients who were distressed about their weight. Clinicians were also significantly more likely to anticipate delaying weighing the patient when patients were distressed about their weight compared to when they were calm. Patient level of distress did not significantly impact clinicians' intended likelihood of relying on self-reported weight, judging weight by eye, or asking someone else to weigh the patient.

These findings support the second hypothesis, as clinicians' intended weighing behaviours were more compliant with guidelines when patient's distress levels were low than when they were distressed about their weight. Despite the trend in scores being in the anticipated direction, the lack of significant differences between the two 'distress' conditions means that hypothesis 3 was not supported.

Table 2

Comparing clinicians' intended likelihood of engaging in different weighing behaviours for patients diagnosed with anorexia nervosa and bulimia nervosa who are experiencing different levels of distress

Weighing Behaviour	Patient Diagnosis	Level of Patient Distress						Friedman ANOVA		Multiple comparisons test (Wilcoxon; $p<.05$ one tailed)	*r*
		Calm		General Distress		Weight Distress		X^2	*p* (2-tailed)		
		M	*SD*	*M*	*SD*	*M*	*SD*				
Weigh patient	Anorexia	94.19	(1.69)	93.38	(1.70)	91.22	(1.83)	14.49	.001	*NS*	-
	Bulimia	93.92	(1.58)	90.14	(2.14)	86.89	(1.65)	29.78	.001	C>WD	.23
Open weigh	Anorexia	95.42	(1.06)	90.97	(1.91)	89.86	(1.65)	31.71	.001	C>WD	.26
	Bulimia	94.25	(1.59)	93.61	(1.31)	90.00	(1.68)	29.12	.001	C>WD	.24
Rely on self-reported weight	Anorexia	6.76	(1.90)	5.14	(1.58)	6.08	(1.69)	0.84	*ns*	-	-
	Bulimia	8.08	(2.09)	8.51	(2.07)	9.86	(2.22)	1.72	*ns*	-	-
Judge weight by eye	Anorexia	12.16	(2.50)	11.35	(2.36)	11.35	(2.33)	1.20	*ns*	-	-
	Bulimia	10.41	(2.22)	10.81	(2.25)	11.76	(2.36)	5.77	*ns*	-	-
Delay weighing	Anorexia	7.95	(2.03)	9.86	(2.15)	13.38	(2.38)	21.42	.001	C<WD	-.21
	Bulimia	7.03	(1.80)	12.88	(2.40)	15.54	(2.43)	22.89	.001	C<WD	-.22
Ask someone else to weigh	Anorexia	5.95	(1.99)	5.68	(1.99)	6.22	(1.94)	3.00	*ns*	-	-
	Bulimia	4.59	(1.65)	4.86	(1.67)	5.54	(1.73)	6.62	*ns*	-	-

Note: C = Calm, GD = General Distress, WD = Weight Distress. *ns* = non significant. Significance level set at $p < .01$.

Hypothesis 4: Clinician characteristics (age, gender, experience, discipline, anxiety, beliefs about open weighing, 'broken leg exceptions', and firm empathy) will be associated with their intended weighing behaviours.

Clinician demographics.

Gender. Mann-Whitney tests were conducted separately for anorexia nervosa and bulimia nervosa patient vignettes, to compare clinicians' intended weighing behaviours by clinician gender (Tables 3 and 4). They demonstrate that clinicians' intended weighing behaviours did not significantly differ according to their gender, which does not support hypothesis 4.

Profession. Kruskal-Wallis tests were conducted separately for anorexia nervosa and bulimia nervosa patient vignettes, comparing clinician's intended weighing behaviours between three professional groupings (Tables 5 and 6). The findings fail to support hypothesis 4, as clinicians' intended weighing behaviours did not significantly differ according to their profession.

Age and years of experience. Spearman's correlation coefficients were conducted separately for anorexia nervosa and bulimia nervosa patients' vignettes, to assess the relationships between clinicians' intended weighing behaviours and their age and years of experience using CBT for eating disorder patients (Table 7). The results do not support hypothesis 4, as clinicians' age and years of experience were not significantly associated with their intended weighing behaviours.

Table 3

Comparison between male and female clinicians' intended likelihood of engaging in different weighing behaviours for patients diagnosed with anorexia nervosa who are experiencing different levels of distress

Weighing Behaviour	Patient's Level of Distress	Females		Males		Mann Whitney Test		
		M	*SD*	*M*	*SD*	*U*	*z*	*p*
Weigh patient	Calm	93.85	(1.78)	98.75	(0.35)	225.50	-0.83	.41
	General distress	93.54	(1.78)	92.50	(1.04)	3.15	1.20	.23
	Weight Distress	92.15	(1.64)	83.75	(3.07)	307.00	0.98	.33
Open weigh	Calm	95.24	(1.12)	96.25	(0.52)	273.50	0.50	.61
	General Distress	90.95	(2.01)	90.00	(1.07)	318.00	1.40	.16
	Weight Distress	89.38	(1.74)	92.50	(0.71)	271.50	0.30	.76
Rely on self-reported weight	Calm	7.38	(2.00)	2.50	(0.71)	270.50	0.30	.77
	General Distress	5.54	(1.67)	2.50	(0.71)	268.50	0.24	.81
	Weight Distress	6.62	(1.79)	2.50	(0.71)	270.50	0.30	.77
Judge weight by eye	Calm	13.38	(2.64)	3.75	(0.52)	261.50	0.30	.97
	General Distress	12.31	(2.49)	5.00	(0.76)	254.00	-0.14	.89
	Weight Distress	11.54	(2.35)	11.25	(2.42)	248.00	-0.26	.79
Delay weighing	Calm	7.66	(2.00)	2.50	(0.71)	286.50	0.74	.46
	General Distress	7.69	(1.90)	18.75	(2.59)	174.50	-1.80	.07
	Weight Distress	11.08	(2.09)	23.75	(3.34)	191.00	-1.36	.17
Ask someone else to weigh	Calm	6.77	(2.11)	0.00	(0.00)	292.00	1.04	.30
	General Distress	6.31	(2.12)	1.25	(0.35)	258.50	-0.05	.96
	Weight Distress	6.92	(2.06)	1.25	(0.35)	271.50	0.33	.74

Note: significance level set at $p < .01$ (2-tailed).

Table 4

Comparison between male and female clinicians' intended likelihood of engaging in different weighing behaviours for patients diagnosed with bulimia nervosa who are experiencing different levels of distress.

Weighing Behaviour	Patient's Level of Distress	Females		Males		Mann Whitney Test		
		M	*SD*	*M*	*SD*	*U*	*z*	*p*
Weigh patient	Calm	93.38	(1.67)	98.75	(0.35)	228.50	-0.77	.44
	General Distress	90.31	(2.21)	90.00	(1.69)	298.00	0.80	.43
	Weight Distress	8.79	(2.22)	78.75	(3.44)	290.00	0.59	.55
Open weigh	Calm	93.75	(1.69)	97.50	(0.35)	252.00	-0.10	.92
	General Distress	93.81	(1.37)	91.25	(0.83)	330.50	1.69	.09
	Weight Distress	89.84	(1.73)	90.00	(1.31)	289.50	0.67	.51
Rely on self-reported weight	Calm	8.75	(2.19)	3.75	(1.06)	268.00	0.34	.73
	General Distress	8.92	(2.12)	6.25	(1.77)	276.00	0.42	.67
	Weight Distress	10.77	(2.33)	3.75	(1.06)	287.50	0.69	.49
Judge weight by eye	Calm	11.38	(2.34)	3.75	(0.74)	273.00	0.30	.76
	General Distress	11.38	(2.32)	7.50	(1.75)	265.00	0.12	.91
	Weight Distress	12.46	(2.48)	7.50	(1.17)	250.50	-0.21	.83
Delay weighing	Calm	6.31	(1.64)	3.75	(1.06)	279.00	0.49	.63
	General Distress	10.47	(2.07)	25.00	(3.74)	192.50	-1.28	.20
	Weight Distress	13.23	(2.04)	26.25	(3.96)	221.50	-0.73	.47
Ask someone else to weigh	Calm	1.52	(1.75)	0.00	(0.00)	288.00	0.97	.33
	General Distress	5.38	(1.78)	1.25	(0.35)	258.00	-0.05	.96
	Weight Distress	6.15	(1.83)	1.25	(0.35)	267.00	0.21	.84

Note: significance level set at $p < .01$ (2-tailed).

Table 5

Comparison between Psychologists', CBT therapists' and other professionals' intended likelihood of engaging in different weighing behaviours for patients diagnosed with anorexia nervosa who are experiencing different levels of distress

Weighing Behaviour	Patient's Level of Distress	Psychologists		CBT Therapists		Other		Kruskal Wallis Test	
		M	*SD*	*M*	*SD*	*M*	*SD*	*H*	*p*
Weigh patient	Calm	98.06	(0.47)	90.83	(1.44)	90.38	(2.58)	5.33	.07
	General Distress	97.22	(0.66)	86.67	(2.81)	91.15	(1.97)	4.98	.08
	Weight Distress	94.72	(1.13)	80.00	(2.86)	91.54	(1.89)	4.62	.10
Open weigh	Calm	96.67	(0.93)	0.93	(0.93)	93.60	(1.29)	2.08	.35
	General Distress	92.22	(1.79)	0.93	(0.93)	87.20	(2.34)	2.36	.31
	Weight Distress	91.39	(1.69)	1.44	(1.44)	87.20	(1.70)	1.61	.45
Rely on self-reported weight	Calm	3.61	(1.44)	8.33	(1.95)	10.38	(2.37)	2.61	.27
	General Distress	2.50	(0.69)	0.00	(0.00)	11.15	(2.45)	3.81	.15
	Weight Distress	3.06	(0.98)	4.17	(1.44)	11.15	(2.39)	2.50	.29
Judge weight by eye	Calm	6.39	(1.85)	2.00	(3.33)	16.54	(2.76)	4.86	.09
	General Distress	6.94	(1.95)	20.83	(3.42)	13.08	(2.24)	3.65	.16
	Weight Distress	7.50	(1.99)	20.00	(3.33)	12.69	(2.20)	3.28	.19
Delay weighing	Calm	3.61	(1.22)	13.64	(2.11)	11.54	(2.72)	5.85	.05
	General Distress	6.39	(1.68)	13.33	(2.81)	13.08	(2.40)	3.10	.21
	Weight Distress	11.11	(2.34)	20.00	(2.86)	13.46	(2.24)	2.09	.35
Ask someone else to weigh	Calm	6.11	(2.14)	4.17	(1.44)	6.54	(2.04)	0.09	.95
	General Distress	6.67	(2.28)	0.00	(0.00)	6.92	(2.04)	1.94	.38
	Weight Distress	6.94	(2.15)	0.00	(0.00)	8.08	(2.08)	2.55	.28

Note: significance level set at $p < .01$ (2-tailed).

Table 6.

Comparison between Psychologists', CBT therapists' and other professionals' intended likelihood of engaging in different weighing behaviours for patients diagnosed with bulimia nervosa who are experiencing different levels of distress

Weighing Behaviour	Patient's Level of Distress	Psychologists		CBT Therapists		Other		Kruskal Wallis Test	
		M	*SD*	*M*	*SD*	*M*	*SD*	*H*	*p*
Weigh patient	Calm	96.67	(0.86)	92.50	(1.48)	90.77	(2.24)	1.12	.58
	General Distress	93.61	(1.62)	83.33	(2.99)	88.46	(2.29)	3.32	.19
	Weight Distress	90.83	(1.90)	73.33	(3.28)	87.69	(2.29)	3.61	.17
Open weigh	Calm	93.89	(1.95)	97.50	(0.62)	93.20	(1.35)	2.79	.25
	General Distress	95.28	(1.23)	96.36	(0.67)	90.00	(1.58)	3.57	.17
	Weight Distress	91.11	(1.77)	92.50	(1.48)	87.20	(1.65)	3.63	.16
Rely on self-reported weight	Calm	2.50	(0.87)	11.82	(2.71)	14.23	(2.76)	3.28	.19
	General Distress	4.17	(1.36)	9.17	(1.93)	14.23	(2.76)	2.31	.32
	Weight Distress	4.72	(1.46)	10.83	(2.61)	16.54	(2.76)	3.08	.22
Judge weight by eye	Calm	6.94	(2.00)	18.33	(2.91)	11.54	(2.13)	2.52	.28
	General Distress	7.50	(2.06)	18.33	(2.92)	11.92	(2.15)	2.41	.30
	Weight Distress	8.89	(2.17)	20.00	(3.33)	11.92	(2.08)	1.48	.48
Delay weighing	Calm	5.56	(1.54)	5.83	(1.44)	9.62	(2.25)	0.14	.94
	General Distress	8.06	(1.79)	20.91	(3.24)	16.15	(2.04)	2.54	.28
	Weight Distress	11.67	(2.18)	26.67	(3.28)	15.77	(2.23)	2.54	.28
Ask someone else to weigh	Calm	4.72	(1.61)	0.00	(0.00)	6.54	(2.04)	1.48	.48
	General Distress	5.00	(1.66)	0.00	(0.00)	6.29	(2.04)	1.99	.37
	Weight Distress	5.56	(1.73)	0.00	(0.00)	8.08	(2.08)	2.57	.28

Note: significance level set at $p < .01$ (2-tailed).

Table 7

Correlations of clinicians' age and years of experience using CBT for eating disorders with their intended likelihood of engaging in different weighing behaviours for patients diagnosed with anorexia nervosa and bulimia nervosa who are experiencing different levels of distress

Weighing Behaviour	Patient Level of Distress	Correlation Coefficient (r_s)			
		Anorexia nervosa		Bulimia nervosa	
		Clinician Age	Clinician Experience	Clinician Age	Clinician Experience
Weigh patient	Calm	-.19	.00	-.18	-.11
	General Distress	-.24	-.11	-.29	-.18
	Weight Distress	-.20	-.09	-.33	-.24
Open weigh	Calm	-.17	-.04	-.18	-.03
	General Distress	-.18	.13	-.12	.04
	Weight Distress	-.08	.14	-.05	.13
Rely on self-reported weight	Calm	.19	.25	.19	.13
	General Distress	.08	.15	.20	.17
	Weight Distress	.10	.15	.16	.25
Judge weight by eye	Calm	.10	-.09	.18	.00
	General Distress	.14	-.03	.21	.02
	Weight Distress	.20	-.01	.22	.04
Delay weighing	Calm	.21	.00	.12	.15
	General Distress	.25	.12	.26	.08
	Weight Distress	.15	.13	.29	.27
Ask someone else to weigh	Calm	.07	.09	.01	.02
	General Distress	.06	-.03	.06	-.03
	Weight Distress	.01	-.05	.02	-.06

Note: No associations were significant at the adjusted alpha of $p < .001$.

Clinician's cognitions and emotions.

Beliefs about open weighing. Spearman's correlation coefficients were conducted separately for anorexia nervosa and bulimia nervosa patient vignettes, to assess the relationships between clinicians' intended weighing behaviours and their beliefs about open weighing (Tables 8 and 9). These findings are interpreted in terms of the broad patterns observed. Isolated statistically significant results are not considered, due to the large number of analyses and possibility of type I error.

Negative beliefs. The belief that 'someone else should weigh the patient' was not associated with intended weighing behaviours for anorexia nervosa patients. However, for bulimia nervosa patients, clinicians who more strongly endorsed this belief were less likely to intend to weigh or open weigh, and were more likely to intend to delay weighing or ask someone else to weigh if the patient was distressed.

For anorexia nervosa patients, clinicians who more strongly endorsed the belief that 'open weighing damages the therapeutic relationship' were less likely to intend to open weigh and more likely to intend to judge the patients weight by eye, regardless of level of distress. For bulimia nervosa patients they were less likely to weigh or open weigh and were more likely to rely on the patient's self-reported weight, judge their weight by eye, and delay weighing, regardless of the patient's distress level.

Clinicians who more strongly endorsed the belief that 'open weighing is harmful to the patient' were less likely to intend to openly weigh and were more likely to intend to judge the patient's weight by eye, regardless of patient's level of distress or diagnosis.

The belief that 'open weighing makes the patient angry' was not significantly associated with clinicians' anticipated weighing behaviours.

Positive beliefs. Increased belief that 'open weighing reduces patient anxiety' was associated with increased likelihood of intending to weigh and openly weigh the patient and decreased likelihood of intending to judge patient weight by eye, for both

anorexia nervosa and bulimia nervosa patients, with some exceptions for patients who were distressed about their weight. Endorsing this belief was also associated with reduced likelihood of intending to delay weighing the patient, but only for bulimia nervosa patients.

Finally, clinicians who more strongly endorsed the belief that ‘open weighing improves patients' understanding’ were more likely to anticipate weighing and openly weighing patients, and were less likely to anticipate relying on patient's self-reported weight, judging patient weigh by eye or delaying weighing, regardless of patient diagnosis or level of distress.

These findings support hypothesis 4, as clinicians' beliefs about open weighing were associated with their intended weighing behaviours. Overall, negative beliefs about open weighing were associated with less guideline-compliant weighing behaviours, while positive beliefs were associated with more guideline-compliant weighing behaviours.

Anxiety. Tables 10 and 11 show the results of Spearman's correlation coefficients, conducted separately for anorexia nervosa and bulimia nervosa patient vignettes. These correlations test hypothesis 4 by assessing the relationships of clinicians' prospective and inhibitory anxiety IUS subscale scores, BLES-OW total scores and firmness and empathy FEQ subscale scores with their intended weighing behaviours. Again, these tables are interpreted in terms of the broad patterns observed. Isolated statistically significant results are not considered, due to the large number of analyses and possibility of type I error. Clinician prospective anxiety and inhibitory anxiety were not associated with their weighing behaviours, which does not support hypothesis 4.

Broken leg exceptions. Tables 10 and 11 indicate that clinicians who endorse more broken leg exceptions are less likely to anticipate weighing or openly weighing

the patient and are more likely to anticipate relying on the patient's self-reported weight, judging patient weight by eye or delaying weighing, regardless of patient diagnosis or level of distress. This supports hypothesis 4, as clinicians' broken leg exceptions were associated with their intended weighing behaviours.

Firm Empathy. Tables 10 and 11 indicate that clinician's firmness and empathy were not significantly associated with their intended weighing behaviours, which does not support hypothesis 4.

Overall the results partially support hypothesis 4. Some clinician characteristics (beliefs about open weighing and 'broken leg exceptions') were associated with their intended weighing behaviours. However, others (age, gender, experience, discipline, anxiety, and firm empathy) were not.

Table 8

Correlations between clinicians' beliefs about open weighing and their intended likelihood of engaging in different weighing behaviours for patients diagnosed with anorexia nervosa who are experiencing different levels of distress

Weighing Behaviour	Patient Level of Distress	Clinician Beliefs about Open Weighing Correlation Coefficient (r_s)					
		Negative Beliefs				Positive Beliefs	
		Someone else should weigh the patient	OW damages the therapeutic relationship	OW is harmful to the patient	OW makes patients angry	OW reduces patient anxiety	OW improves patient understanding
Weigh patient	Calm	-.04	-.25	-.14	-.02	.35*	.42**
	General Distress	-.22	-.16	-.23	-.20	.31*	.36*
	Weight Distress	-.20	-.20	-.21	-.03	.30	.34*
Open weigh	Calm	-.26	-.34*	-.42**	-.28	.35*	.41**
	General Distress	-.33	-.42**	-.37*	-.21	.37*	.45**
	Weight Distress	-.33	-.35*	-.37*	-.08	.33*	.33*
Rely on self-reported weight	Calm	.26	.26	.21	.10	-.18	-.32*
	General Distress	.22	.19	.21	.30*	-.12	-.32*
	Weight Distress	.21	.27	.22	.11	-.23	-.32*
Judge weight by eye	Calm	.25	.40**	.38*	.07	-.43**	-.52**
	General Distress	.29	.36*	.42**	.13	-.41**	-.49**
	Weight Distress	.29	.33*	.40**	.09	-.36*	-.42**
Delay weighing	Calm	.06	.25	.14	.03	-.34*	-.43**
	General Distress	.26	.13	.19	.20	-.27	-.39*
	Weight Distress	.21	.21	.19	.14	-.22	-.42**
Ask someone else to weigh	Calm	.19	.16	.14	.11	-.02	-.10
	General Distress	.24	.27	.25	.18	-.11	-.20
	Weight Distress	.28	.33*	.28	.16	-.11	-.25

Note: OW= Open weighing. Adjusted significance level set at $p < .01$ (2-tailed). * $p < .01$, ** $p < .001$.

Table 9

Correlations between clinicians' beliefs about open weighing and their intended likelihood of engaging in different weighing behaviours for patients diagnosed with bulimia nervosa who are experiencing different levels of distress

Weighing Behaviour	Patient's Level of Distress	Clinician Beliefs about Open Weighing Correlation Coefficient (r_s)					
		Negative Beliefs				Positive Beliefs	
		Someone else should weigh the patient	OW damages the therapeutic relationship	OW is harmful to the patient	OW makes patients angry	OW reduces anxiety	OW improves patient understanding
Weigh patient	Calm	-.21	-.36*	-.32*	-.24	.41**	.47**
	General Distress	-.36*	-.38*	-.30	-.24	.48**	.49**
	Weight Distress	-.38*	-.29	-.21	-.12	.37*	.42**
Open weigh	Calm	-.27	-.36*	-.39*	-.21	.30*	.34*
	General Distress	-.42**	-.45**	-.39*	-.32*	.34*	.38*
	Weight Distress	-.39*	-.45**	-.42**	-.22	.29	.35*
Rely on self-reported weight	Calm	.24	.34*	.29	.18	-.29	-.41**
	General Distress	.27	.33*	.28	.15	-.26	-.40**
	Weight Distress	.25	.26	.27	.09	-.22	-.31*
Judge weight by eye	Calm	.29	.33*	.40**	.13	-.40**	-.49**
	General Distress	.28	.32*	.39*	.13	-.40**	-.48**
	Weight Distress	.29	.32*	.38*	.13	-.38*	-.47**
Delay weighing	Calm	.19	.31*	.28	.23	-.40**	-.45**
	General Distress	.36*	.49**	.31*	.24	-.52**	-.58**
	Weight Distress	.40**	.30*	.21	.13	-.27	-.43**
Ask someone else to weigh	Calm	.19	.20	.18	.19	-.09	-.15
	General Distress	.25	.28	.25	.19	-.12	-.21
	Weight Distress	.33*	.39*	.34*	.20	-.16	-.32*

Note: OW= Open weighing. Adjusted significance level set at $P < .01$ (2-tailed). * $P < .01$, ** $P < .001$.

Table 10

Correlations between clinicians' levels of intolerance of uncertainty, firm empathy and likelihood of making broken leg exceptions, and their intended likelihood of engaging in different weighing behaviours for patients diagnosed with anorexia nervosa who are experiencing different levels of distress

Weighing Behaviour	Patient Level of Distress	Clinician Characteristics Correlation Coefficients (r_s)				
		Intolerance of uncertainty		Broken Leg Exceptions	Firm Empathy	
		Prospective Anxiety	Inhibitory Anxiety		Empathy	Firmness
Weigh patient	Calm	-.09	-.11	-.39*	.07	-.01
	General Distress	-.01	-.10	-.45**	.09	.03
	Weight Distress	-.11	-.18	-.41**	.12	.03
Open weigh	Calm	.02	-.10	.46**	.07	.02
	General Distress	-.04	-.20	-.44**	.19	.15
	Weight Distress	-.04	-.14	-.39*	.13	.08
Rely on self-reported weight	Calm	.13	.08	.42**	-.20	-.13
	General Distress	.14	.15	.45**	-.15	-.14
	Weight Distress	.11	.15	.41**	-.19	-.13
Judge weight by eye	Calm	.04	.07	.38*	.09	.13
	General Distress	.09	.11	.40**	.09	.12
	Weight Distress	-.00	.05	.40**	.00	.04
Delay weighing	Calm	.10	.11	.40**	-.09	-.01
	General Distress	.03	.08	.42**	-.14	-.09
	Weight Distress	.09	.16	.42**	-.16	-.09
Ask someone else to weigh	Calm	.06	.05	.24	-.10	-.04
	General Distress	.02	.05	.28	-.08	-.05
	Weight Distress	-.01	.12	.34*	-.15	-.11

Note: Adjusted significance level set at $p < .01$ (2-tailed). * $p < .01$, ** $p <. 001$.

Table 11

Correlations between clinicians' levels of intolerance of uncertainty, firm empathy and likelihood of making broken leg exceptions and their intended likelihood of engaging in different weighing behaviours for patients diagnosed with bulimia nervosa who are experiencing different levels of distress

Weighing Behaviour	Patient's Level of Distress	Clinician Characteristics Correlation Coefficients (r_s)				
		Intolerance of Uncertainty		Broken Leg Exceptions	Firm Empathy	
		Prospective Anxiety	Inhibitory anxiety		Empathy	Firmness
Weigh patient	Calm	-.12	-.21	-.56**	.17	.10
	General Distress	-.04	-.22	-.43**	.20	.13
	Weight Distress	-.02	-.27	-.38*	.31*	.23
Open weigh	Calm	.01	-.12	-.34*	.05	.02
	General Distress	-.04	-.26	-.46**	.31	.28
	Weight Distress	-.02	-.17	.46**	.20	.17
Rely on self-reported weight	Calm	.22	.18	.46**	-.14	-.10
	General Distress	.20	.15	.45**	-.18	-.14
	Weight Distress	.04	.05	.45**	-.21	-.15
Judge weight by eye	Calm	.08	.11	.41**	.05	.10
	General Distress	.06	.11	.40**	.05	.11
	Weight Distress	.02	.10	.46**	-.02	.03
Delay weighing	Calm	.05	.15	.50**	-.16	-.08
	General Distress	.08	.24	.46**	-.15	-.60
	Weight Distress	.12	.33*	.46**	.41**	.35*
Ask someone else to weigh	Calm	.08	.10	.23	-.07	-.03
	General Distress	.01	.05	.29	-.09	-.05
	Weight Distress	.08	.09	.36*	-.18	-.15

Note: Adjusted significance level set at $p < .01$ (2-tailed). * $p < .01$. ** $p < .001$.

Discussion

This experimental study examined whether patient and clinician characteristics impact on clinicians' anticipated weighing behaviours. This discussion will summarise the main results and relate them to the existing evidence base and current theory. Limitations of the study, recommendations for further research and clinical implications will be considered.

Summary of results

Most clinicians stated that they would weigh their patients more often than not - a finding that is at odds with the existing evidence regarding routine clinical practice (e.g., Cowdrey & Waller, 2015; Forbush et al., 2015; Mulkens et al., under consideration; Waller et al., 2012). It is likely that this difference is due to the demand characteristics of this specific study. However, the area of interest in this study is the difference in clinicians' anticipated weighing behaviours under different experimental conditions, rather than overall levels of those behaviours.

Patient characteristics (hypotheses 1-3). Patient characteristics impacted on clinicians' anticipated weighing behaviours. Clinicians reported being more likely to weigh patients diagnosed with anorexia nervosa, but were more likely to rely on self-reported weight if the patient was diagnosed with bulimia nervosa, supporting hypothesis 1. Clinicians' intended weighing behaviours were more compliant with guidelines when patients' were calm rather than distressed about their weight, supporting hypothesis 2. However, hypothesis 3 was not supported, as there were no significant differences in clinicians' anticipated weighing behaviours between different forms of high distress (general vs. weight-related).

Clinician characteristics (hypothesis 4). Hypothesis 4 predicted that clinician characteristics would be associated with their intended weighing behaviours, and this was partially supported. Several clinician characteristics (age, gender, years of

experience, professional discipline, anxiety, and firm empathy) were not related to their weighing behaviours. However, clinicians' beliefs about open weighing and tendency to make 'broken leg exceptions' were associated with their intended weighing behaviours, as outlined below.

Overall, negative beliefs about open weighing (e.g., 'someone else should weigh the patient'; 'it damages the therapeutic relationship'; 'it is harmful to the patient') were associated with fewer guideline-compliant weighing behaviours. Conversely, positive beliefs ('open weighing reduces patient anxiety'; 'it improves patient understanding') were associated with more guideline-compliant weighing behaviours. However, some of those links between beliefs about open weighing and intended behaviour in clinic appear to be moderated by patient characteristics (diagnosis and level of distress).

Clinicians who endorsed more broken leg exceptions were more likely to use non-compliant weighing behaviours, regardless of patient characteristics.

Relationship to the existing evidence

Broadly, the findings of this study contribute to the growing evidence base on 'therapist drift' (Waller, 2009) in psychotherapy. They support previous findings that clinician and patient characteristics impact clinicians' use of CBT protocol prescribed behavioural techniques (Waller & Turner, 2016), such as weighing the patient.

The importance of clinicians' beliefs about open weighing in this study, supports existing literature on the impact of therapist beliefs on whether they will use exposure-based interventions in other disorders (e.g., in PTSD; Becker, Zayfert & Anderson, 2004; van Minnen, Hendricks & Olff, 2010), and how they will do so (e.g., in anxiety disorders and panic; Deacon, Farrell et al., 2013; Deacon, Lickel, Farrell, Kemp & Hipo, 2013; Farrell, Deacon, Kemp, Dixon & Sy, 2013). The common theme across this and previous studies is that many clinicians reduce the demands of therapy on their patients, based on the therapist's negative beliefs rather than the evidence about what

works. However, this study has also demonstrated that positive clinician beliefs about a CBT technique (in this case, open weighing) can result in greater use of that technique.

This study also supports the conclusion that clinicians make 'broken leg exceptions', excluding patients from some CBT techniques based on the clinician's beliefs without justifiable reasons or evidence (e.g., Meyer et al., 2014).

Other findings of this study are not compatible with the wider evidence base. Clinicians' anxiety was not related to their weighing behaviours, despite its link to the use of other behavioural interventions (e.g., Levita, Gonzales Salas Duhne, Girling, & Waller, 2016; Meyer et al., 2014; Waller et al., 2012; Waller & Turner, 2016). Nor was clinicians' firm empathy related to their weighing behaviours, despite some evidence linking empathy to cautious exposure delivery (Farrell, Deacon, Kemp et al., 2013) and clinical observation that 'firm empathy' is needed to deliver CBT effectively (Wilson et al., 1997). Finally, clinician demographic characteristics were not related to their weighing behaviours, adding to the inconclusive evidence around these factors as predictors of clinicians' use of CBT techniques (e.g., Brown et al., 2014; Forbush et al., 2015; van Minnen et al., 2010; Waller et al., 2012).

Considering patient characteristics, this study found no significant differences in clinicians' intentions to openly weigh patients across diagnoses, unlike Forbush et al. (2015). However, there were differences in weighing behaviours across diagnoses. Clinicians were more likely to weigh anorexia nervosa patients, albeit blind, which might be a compromise between not wanting to distress the patient but recognising their greater risk of malnutrition (Beat, 2015).

Whilst Forbush et al. (2015) found clinicians' weighing behaviours were less guideline-compliant when patients were emotionally distressed, this study suggests it may not be the *intensity* of patients' distress that impacts clinicians' behaviour, but the *specific nature* of distress being about their weight. Previous work in panic disorder has

found that clinicians are less likely to use exposure when patients are reluctant to engage in it (Kazantzis, Ford, Paganini, Dattilio & Farchione, 2017). It is possible that clinicians perceived patient weight-related distress as resistance to being weighed, causing the clinician to back off from behavioural techniques (Waller et al., 2012).

Relationship to Theory

Many theories posit that attitudes, which contain affective and cognitive components, drive behaviour (e.g., Ajzen, 1991). Similarly, therapist drift is thought to occur when clinician's beliefs and emotions lead them to develop safety behaviours, including avoidance of behavioural techniques **(**Waller & Turner, 2016). However, of the clinician variables assessed in this study, only beliefs were related to clinicians weighing behaviours, while anxiety was not. This suggests that cognitive components might be more salient than affective components in influencing clinicians' use of weighing. Alternatively, it may be that state anxiety and trait anxiety act differently, as only trait anxiety was assessed. There may have been state anxiety changes in response to specific vignette factors (e.g., patient weight-related distress) that were not assessed.

This study suggests that clinicians are less likely to anticipate weighing or openly weighing patients if they believe open weighing would do harm or be contraindicated. Instead clinicians are more likely to engage in safety behaviours and avoid weighing to reduce short-term distress. Conversely, findings suggest that clinicians are more likely to anticipate weighing and openly weighing patients if they believe it would do good in the long-term. Patient diagnosis and level of distress may moderate this relationship. It may be that where clinicians perceive greater health risks of not weighing, this overrides concerns of short-term distress. Conversely patients exhibiting weight related distress may increase the salience of short-term harm to clinicians, making them more likely to avoid weighing.

Patient diagnosis and distress are inherent to their disorder, so it is clinician beliefs that must change to reduce these safety behaviours and therapist drift. However, clinicians' negative beliefs may persist as they avoid exposure and then misattribute the lack of short-term negative outcomes to this avoidance. In avoiding exposure, they lack the opportunity to receive disconfirming evidence that exposure has positive long-term outcomes for the patient (Deacon, Lickel et al., 2013).

Limitations

This study has several limitations. The convenience, snowball sampling method might have resulted in self-selection bias and made the findings less generalisable. However, given the need to recruit volunteers, it would have been difficult to obtain a randomised sample. A strength of the recruitment strategy was that the sample was international and included a range of professional disciplines, making it more representative.

In addition, the study was clearly looking at adherence to a guideline-recommended behaviour and relied on clinicians' self-report. Therefore, there is a risk of bias and socially desirable responding. In order to reduce the impact of such biases, the study focused on differences between responses to the vignette conditions, rather than clinicians' overall rating of their behaviour.

It is also acknowledged that the design has limited ecological validity. It might have been improved by observing clinicians' weighing behaviours in real-life clinical settings. However, this experimental design was appropriate for this study, given the scarcity and correlational nature of the existing research into clinician weighing behaviours.

The questionnaires used in this study might be critiqued, as some measures were adapted or created. This approach improved their relevance, but meant that they were not previously validated. However, the BLES-OW showed good internal consistency

and might be useful in further research into clinician weighing behaviours. The study could have used the pre-existing and validated Therapist Beliefs about Exposure Scale (TBES; Deacon, Farrell et al., 2013). However, that scale contains items that are not applicable to open weighing. Instead, the questions regarding beliefs about open weighing were constructed based on clinical experience and the existing literature, and were analysed as individual items.

In terms of the analysis, data violated the assumptions of normality, necessitating the use of non-parametric statistics. The original power calculation was based on the planned parametric statistics, but the power of non-parametric tests is difficult to calculate, as the type 1 error rate is unknown when the sampling distribution is not normally distributed (Field, 2013). However, recruitment exceeded expectations and there was no pattern of findings that merely approached significance, suggesting this study was sufficiently powered.

It is important to acknowledge that where multiple tests were conducted the significance level was generally adjusted to $P < .01$ to reduce the likelihood of incorrectly rejecting the null hypothesis. Bonferroni corrections were not used as this calculation was deemed too conservative, given the hypothesis-driven nature of the study. However no corrections or adjustments were made for the primary analysis assessing the impact of patient diagnosis on clinicians' weighing behaviors, which involved fewer, but still multiple, tests of significance. This increases the possibility of Type 1 error and limits the interpretation of these findings.

Future Research

Further research should seek to address the methodological limitations outlined, using a more robust sampling method. Ecological validity should also be addressed with further research observing clinicians' weighing behaviours in real-life clinical settings with real patients. In terms of focusing that real-world research, this study

suggests that it is important to consider a range of clinician weighing behaviours, and particularly to explore the impact of clinician beliefs on their weighing behaviours.

Future studies of clinician weighing behaviours should further explore the impact of patients' specific weight-related distress over general distress. Do clinicians interpret patients' weight-related distress as reluctance to engage in weighing, even though such reluctance has not been stated by the patient? Answering this question could help to clarify the role of beliefs about patient fragility (Meehl, 1973) or fear of patient resistance (Waller et al., 2012). Research should also consider whether clinician fears about doing harm underlie the impact of patient distress on clinician weighing behaviours, perhaps by increasing the salience of negative beliefs about short-term-harm from open weighing. Assessing clinicians' state anxiety when faced with these scenarios would also assist in exploring this connection.

In addition to observing clinicians in vivo, patients' experiences of the therapy they receive should also be studied (e.g., Cowdrey & Waller, 2015), to triangulate the existing research on therapists' accounts of their past or anticipated weighing behaviours. Considering patient outcomes in relation to clinician weighing behaviours is another avenue for further research.

Finally, this study only considered clinicians' using CBT for eating disorders. However, open weighing is also recommended in the other NICE recommended treatments – MANTRA, and SSCM. Further research into clinician weighing behaviours should include clinicians using these approaches.

Clinical Implications

All evidence-based therapies for eating disorders incorporate open weighing (Waller & Mountford, 2015). However, many clinicians either do not weigh their patients at all (Waller et al., 2012) or weigh 'blind' (Forbush et al., 2015). This study suggests patients might be excluded from open weighing simply due to their individual

clinicians' negative beliefs about open weighing and use of broken leg exceptions. Therefore, many clinicians' weighing behaviours and their underlying beliefs need to change.

Critically, when using CBT-ED, clinicians must openly weigh all patients in each session, regardless of the clinician's personal beliefs about open weighing. Clinicians should monitor their practice and acknowledge that if they fail to open weigh then they are making a choice to practice poorly. Clinicians should be prepared to challenge their negative beliefs about open weighing or exposure more broadly by putting it into practice with patients and monitoring outcomes.

It is important that clinicians receive regular supervision, which includes honest discussion of patient outcome measures (including weight). Supervisors should directly address clinicians' own safety behaviours around weighing patients. They should support clinicians to explore the beliefs that underpin their decisions around weighing, and to implement behavioural changes.

Services and commissioners also have a responsibility to ensure that evidence-based practice is being implemented and should mandate and monitor clinicians' use of regular open weighing in CBT-ED. Services should also monitor the positive and negative impacts of open weighing in terms of patient outcomes and patient experiences of treatment. This information should be available to clinicians, so they can clearly see the impact of open weighing in their service.

Finally, services should also provide clinicians with training, including psycho-education on the guidelines and rationale for open weighing, and the lack of evidence for broken leg exceptions. Training should focus on experiential activities targeting implicit learning, which are more likely to lead to behavioural change (e.g., Deacon, Lickel et al., 2013). Experiential activities might include behavioural experiments to test

out clinicians' negative predictions about open weighing, and role plays exposing clinicians to exposure themselves, so they can learn first-hand that it is not harmful.

Conclusion

This study has demonstrated that in CBT for eating disorders, clinicians' anticipated weighing behaviours might be influenced by their negative beliefs about open weighing and their tendency to make broken leg exceptions. These influences lead the clinicians to bypass this element of treatment guidelines, particularly in the face of patient characteristics such as diagnosis and weight-related distress. Based on the findings of this study, this report has highlighted the potential clinical implications for patients who do not receive evidence-based interventions, and provides recommendations for clinician training and supervision around open weighing. Whilst there are some methodological limitations, this study is the first to use an experimental design to look at patient and clinician characteristics in relation to a wide range of clinician weighing behaviours. As such, it contributes both to the evidence on clinician weighing behaviours and to the wider literature on therapist drift and exposure.

References

Addis, M. E., & Krasnow, A. D. (2000). A national survey of practicing psychologists' attitudes toward psychotherapy treatment manuals. *Journal of Consulting and Clinical Psychology, 68*, 331-339. doi:10.1037/0022-006X.68.2.331

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (5th ed). Arlington, VA: American Psychiatric Publishing.

Azjen, I. (1991). The theory of planned behaviour. *Organizational Behavior and Human Decision Processes, 50*, 179-211. doi:10.1016/0749-5978(91)90020

Beat (2015). *The costs of eating disorders Social, health and economic impacts*. [online] London: PricewaterhouseCoopers. Retrieved from: http://cdn.basw.co.uk/upload/basw_104500-2.pdf

Becker, C. B., Zayfert, C., & Anderson, E. (2004). A survey of psychologists' attitudes towards and utilization of exposure therapy for PTSD. *Behaviour Research and Therapy*, *42*, 277–292. doi:10.1016/S0005-7967(03)00138-4

Brown, A., Mountford, V., & Waller, G. (2014). Clinician and practice characteristics influencing delivery and outcomes of the early part of outpatient cognitive behavioural therapy for anorexia nervosa. *The Cognitive Behaviour Therapist, 7*, e1-e7. doi:10.1017/S1754470X14000105

Carleton, R. N., Mulvogue, M. K., Thibodeau, M. A., McCabe, R. E., Antony, M. M., & Asmundson, G. J. (2012). Increasingly certain about uncertainty: intolerance of uncertainty across anxiety and depression. *Journal of Anxiety Disorders*, *26*, 468–479. doi:10.1016/j.janxdis.2012.01.011

Carleton, R. N., Norton, M. A., & Asmundson, G. J. G. (2007). Fearing the unknown: A short version of the intolerance of uncertainty scale. *Journal of Anxiety Disorders, 21*, 105-117. doi:10.1016/j.janxdis.2006.03.014

Cowdrey, N. D., & Waller, G. (2015). Are we really delivering evidence-based treatments for eating disorders? How eating-disordered patients describe their experience of cognitive behavioural therapy. *Behaviour Research and Therapy, 75*, 72–77. doi:10.1016/j.brat.2015.10.009

Deacon, B. J., Farrell, N. R., Kemp, J. J., Dixon, L. J., Sy, J. T., Zhang, A. R., & McGrath, P. B. (2013). Assessing therapist reservations about exposure therapy for anxiety disorders: The Therapist Beliefs about Exposure Scale. *Journal of Anxiety Disorders, 27*, 772–780. doi:10.1016/j.janxdis.2013.04.006

Deacon, B. J., Lickel, J. J., Farrell, N. R., Kemp, J. J., & Hipol, L. J. (2013). Therapist perceptions and delivery of interoceptive exposure for panic disorder. *Journal of Anxiety Disorders, 27*, 259–264. doi:10.1016/j.janxdis.2013.02.004

Fairburn, C. G., & Harrison, P. J. (2003). Eating disorders. *Lancet, 361*, 407-416. doi:10.1016/S0140-6736(03)12378-1

Farrell, N. R., Deacon, B. J., Dixon, L. J., & Lickel, J. J. (2013). Theory-based training strategies for modifying practitioner concerns about exposure therapy. *Journal of Anxiety Disorders*, 27, 781-787. doi:10.1016/j.janxdis.2013.09.003

Farrell, N. R., Deacon, B. J., Kemp, J. J., Dixon, L. J., & Sy, J. T. (2013). Do negative beliefs about exposure therapy cause its suboptimal delivery? An experimental investigation. *Journal of Anxiety Disorders, 27*, 763–771. doi:10.1016/j.janxdis.2013.03.007

Faul, F., Erdfelder, E., Lang, A. G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioural, and biomedical sciences. *Behaviour Research Methods, 39*, 175-191. doi:10.3758/BF03193146

Field, A. (2013). *Discovering Statistics using IBM SPSS Statistics* (4th Ed.). London, UK: SAGE Publications Ltd.

Forbush, K. T., Richardson, J., & Bohrer, B. K. (2015). Clinicians' practices regarding blind versus open weighing among patients with eating disorders. *International Journal of Eating Disorders*. *48*, 905-911. doi:1002/eat.22369

Hay, P. (2013). A systematic review of evidence for psychological treatments in eating disorders: 2005–2012. *International Journal of Eating Disorders, 46,* 462–469. doi:10.1002/eat.22103

Kazantzis, N., Ford, C., Paganini, C., Dattilio, F. M., & Farchione, D. (2017). Does patient reluctance towards exposure and psychologists' attitudes about evidence based practice influence treatment recommendations for panic disorder? An experimental investigation. *Journal of Anxiety Disorders, 51*, 55–64. doi:10.1016/j.janxdis.2017.03.001

Levita, L., Gonzalez Salas Duhne, P., Girling, C., & Waller, G. (2016). Facets of clinicians' anxiety and the delivery of cognitive behavioral therapy. *Behaviour Research and Therapy*, *77*, 157–161. doi:10.1016/j.brat.2015.12.015

McAdam Freud, E., & Waller, G. (under consideration). Therapists' traits and their relation to preferred therapeutic orientation: The role of firmness and empathy. *Psychology and Psychotherapy.*

McEvoy, P. M., & Mahoney, A. E. J. (2011). Achieving certainty about the structure of intolerance of uncertainty in a treatment-seeking sample with anxiety and depression. *Journal of Anxiety Disorders, 25,* 112-122. doi:10.1016/j.janxdis.2010.08.010

Meehl, P. E. (1973). Why I do not attend case conferences. In P.E. Meehl (Ed) *Psychodiagnosis: Selected papers* (pp. 225-302). Minneapolis, MN: University of Minnesota Press.

Meyer, J. M., Farrell, N. R., Kemp, J. J., Blakey, S. M., & Deacon, B. J. (2014). Why do clinicians exclude anxious clients from exposure therapy? *Behaviour Research and Therapy, 54,* 49-53. doi:10.1016/j.brat.2014.01.004

Mulkens, S., de Vos, C., de Graff, A., & Waller, G. (under consideration). To deliver or not to deliver Cognitive Behavioral Therapy for Eating Disorders: Replication and extension of our understanding of why therapists fail to do what they should do. *Journal of Consulting and Clinical Psychology.*

National Institute for Health and Care Excellence. (2017). *Eating disorders: recognition and treatment, NG69*. London, UK: National Institute for Health and Care Excellence. Retrieved from: https://www.nice.org.uk/guidance/ng69

Rosenthal, R. (1991). *Meta-analytic procedures for social research* (2nd Ed.). Newbury Park, CA: Sage.

Steinhausen, H. C. (2002). The outcome of anorexia nervosa in the 20th centrury. *American Journal of Psychiatry, 159,* 1284-1293. doi:10.1176/appi.ajp.159.8.1284

Steinhausen, H. C., & Weber, S. (2009). The outcome of bulimia nervosa: Findings from one-quarter century of research. *The American Journal of Psychiatry, 166,* 1331-1341. doi:10.1176/appi.ajp.2009.09040582

Tobin, D. L., Banker, J. D., Weisberg, L., & Bowers, W. (2007) I know what you did last summer (and it was not CBT): a factor analytic model of international psychotherapeutic practice in the eating disorders. *International Journal of Eating Disorders, 40,* 754-757. doi:10.1002/eat.20426

Turner, H., Tatham, M., Lant, M., Mountford, V. A., & Waller, G. (2014). Clinicians' concerns about delivering cognitive-behavioural therapy for eating disorders. *Behaviour Research and Therapy, 57,* 38-42. doi:10.1016/j.brat.2014.04.003

van Minnen, A., Hendriks, L., & Olff, M. (2010). When do trauma experts choose

exposure therapy for PTSD patients? A controlled study of therapist and patient factors. *Behaviour Research and Therapy, 48*, 312–320. doi:10.1016/j.brat.2009.12.003

Waller, G. (2009). Evidence-based treatment and therapist drift. *Behaviour Research and Therapy, 47*, 119-127. doi:10.1016/j.brat.2008.10.018

Waller G., & Mountford, V. (2015). Weighing patients within cognitive-behavioural therapy for eating disorders: how, when and why. *Behaviour Research and Therapy. 70*, 1-10. doi:10.1016/j.brat.2015.04.004

Waller, G., Mountford, A., Tatham, M., Turner, H. Gabriel, C., & Webber, R. (2013). Attitudes towards psychotherapy manuals among clinicians treating eating disorders. *Behaviour Research and Therapy, 51*, 840-844. doi:10.1016/j.brat.2013.10.004

Waller, G., Stringer, H., & Meyer, C. (2012). What cognitive-behavioural techniques do therapists report using when delivering cognitive-behavioural therapy for the eating disorders? *Journal of Consulting and Clinical Psychology, 80,* 171-175. doi:10.1037/a0026559

Waller, G., & Turner, H. (2016). Therapist drift redux: Why well-meaning clinicians fail to deliver evidence-based therapy, and how to get back on track. *Behaviour Research and Therapy, 77,* 129-137. doi:10.1016/j.brat.2015.12.005

Wilson, G. T., Fairburn, C. G., & Agras, W. S. (1997). Cognitive behavioural therapy for bulimia nervosa. In D. M. Garner, & P. E. Garfinkel, (Eds.), *Handbook of Treatment for Eating Disorders* (pp. 67–93). New York, NY: Guilford.

Appendix A

Ethical Approval Letter and Clarification of Title Letter

Downloaded: 28/03/2017
Approved: 15/12/2016

Glenn Waller
Psychology

Dear Glenn

PROJECT TITLE: Clinician weighing behaviours with eating disorder patients: Associations with patient and clinician characteristics
APPLICATION: Reference Number 012253

On behalf of the University ethics reviewers who reviewed your project, I am pleased to inform you that on 15/12/2016 the above-named project was **approved** on ethics grounds, on the basis that you will adhere to the following documentation that you submitted for ethics review:

- University research ethics application form 012253 (dated 06/12/2016).
- Participant information sheet 1025637 version 1 (06/12/2016).
- Participant consent form 1025638 version 1 (06/12/2016).

The following optional amendments were suggested:

Please remove the registrar from the information sheet. This office no longer exists. Please check with Tom Webb as Chair of the ethics committee as to whom should be contacted with regards to complaints. It might be someone in RIS, or Tom, or myself for DClins. We need to modify our advice about this so that this does not become a persistent problem and I would therefore be grateful if you would notify Amrit of the outcome of your discussions so that he can modify exemplar information sheets. Advice only - It would seem a shame not to gather more information about the therapist beyond there characteristics - this is not a must do, but it would seem appropriate to consider their anxiety levels or likliness to follow protocols ect ect ...

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

Yours sincerely

Thomas Webb
Ethics Administrator
Psychology

The University Of Sheffield.

Department Of Psychology.
Clinical Psychology Unit.

Doctor of Clinical Psychology (DClin Psy) Programme
Clinical supervision training and NHS research training & consultancy.

Clinical Psychology Unit
Department of Psychology
University of Sheffield
Floor D, Cathedral Court
1 Vicar Lane
Sheffield
S1 1HD

Amrit Sinha Research Support Officer
Telephone: 0114 2226650
Fax: 0114 2226610
Email: a.sinha@sheffield.ac.uk

04 April 2017

Dear Sir/Madam,
As the Research Support Officer for the Dclin Psychology Program; I would like to confirm that there is a slight discrepancy between the final research title and the title as recorded on the ethical approval letter for Dclin Psychology trainee Amy Daglish.

The title as stated on the ethical approval letter is: “Clinician weighing behaviours with eating disorder patients: **Associations** with patient and clinician characteristics.”

The Actual title of the research study should be: "Clinician weighing behaviours with eating disorder patients: The **role of** patient and clinician characteristics". The correct title is registered on the URMS database.

I've spoken with the trainee and Research Innovation Services (RIS) and have confirmed with them the only change in the title is the word: “**Associations with** patient” which should in fact be : “The **role of** patient”.

I am happy that there is no material change to the trainee’s project warranting an amended ethics letter and the trainee and Research Innovation Services (RIS) have agreed this point.
As a result please find this letter affixed to confirm the ethical approval letter is valid; and to confirm the current title difference does not materially affect the substance of the ethical approval letter.

Best wishes,
Amrit

A. Sinha

Amrit Sinha

Research Support Officer

Department of Psychology
University of Sheffield
Cathedral Court

1 Vicar Lane
Sheffield
S1 2LT

Tel: +44 (0) 114 2226650

a.sinha@sheffield.ac.uk

CPU work days: Monday, Tuesday, Thursday

Appendix B

G*Power Sample Size Analysis Output

[1] *-- Thursday, July 21, 2016 -- 08:20:33*

F tests - ANOVA: Repeated measures, within factors

Analysis:	A priori: Compute required sample size		
Input:	Effect size f	=	0.25
	α err prob	=	0.05
	Power (1−β err prob)	=	0.8
	Number of groups	=	2
	Number of measurements	=	3
	Corr among rep measures	=	0.5
	Nonsphericity correction ϵ	=	1
Output:	Noncentrality parameter λ	=	10.5000000
	Critical F	=	3.1751410
	Numerator df	=	2.0000000
	Denominator df	=	52.0000000
	Total sample size	=	28
	Actual power	=	0.8115602

[2] *-- Thursday, July 21, 2016 -- 08:21:23*

F tests - ANOVA: Repeated measures, within factors

Analysis:	A priori: Compute required sample size		
Input:	Effect size f	=	0.20
	α err prob	=	0.05
	Power (1−β err prob)	=	0.8
	Number of groups	=	2
	Number of measurements	=	3
	Corr among rep measures	=	0.5
	Nonsphericity correction ϵ	=	1
Output:	Noncentrality parameter λ	=	10.0800000
	Critical F	=	3.1107662
	Numerator df	=	2.0000000
	Denominator df	=	80.0000000
	Total sample size	=	42
	Actual power	=	0.8031391

Appendix C

Recruitment Email

Hi

Could I ask a favour? We are doing some research on the decisions that CBT clinicians make when deciding whether/how/when to weigh patients with eating disorders. If you would be willing to consider taking part, it would be much appreciated.

If you are willing, please click on the link below, and you will be taken to an online survey, starting with an information sheet and consent form. This has all been approved by our local ethics committee.

https://sheffieldpsychology.eu.qualtrics.com/jfe/form/SV_8HpV5NDakSxWSSp

If you would pass it on to colleagues, that would be even better.

Many thanks

Glenn Waller and Amy Daglish

Appendix D

Intolerance of Uncertainty Scale- Short Form (IUS-12)

Please circle the number that best corresponds to how much you agree with each item.

	Not at all characteristic of me	A little characteristic of me	Somewhat characteristic of me	Very characteristic of me	Entirely characteristic of me
1. Unforeseen events upset me greatly.	1	2	3	4	5
2. It frustrates me not having all the information I need.	1	2	3	4	5
3. Uncertainty keeps me from living a full life.	1	2	3	4	5
4. One should always look ahead so as to avoid surprises.	1	2	3	4	5
5. A small unforeseen event can spoil everything, even with the best of planning.	1	2	3	4	5
6. When it's time to act, uncertainty paralyses me.	1	2	3	4	5
7. When I am uncertain I can't function very well.	1	2	3	4	5
8. I always want to know what the future has in store for me.	1	2	3	4	5
9. I can't stand being taken by surprise.	1	2	3	4	5
10. The smallest doubt can stop me from acting.	1	2	3	4	5
11. I should be able to organize everything in advance.	1	2	3	4	5
12. I must get away from all uncertain situations.	1	2	3	4	5

Appendix E

Beliefs about Open Weighing Questions

Please indicate how much you agree with each of the following statements about different weighing practices in CBT.

Someone other than the CBT therapist should weigh the patient.

Weighing the patient and telling them their weight damages the therapeutic relationship in CBT.

In CBT sessions, weighing the patient and telling them their weight is harmful to them.

In CBT sessions, weighing the patient and telling them their weight makes them angry or confrontational.

In CBT sessions, weighing the patient and telling them their weight helps to reduce the patient's anxiety about weight.

In CBT sessions, weighing the patient and telling them their weight helps them understand the relationship between food and weight.

For each question, participants were presented with the following options:

Strongly disagree	Disagree	Somewhat disagree	Neither disagree nor agree	Somewhat agree	Agree	Strongly agree
○	○	○	○	○	○	○

Appendix F

Broken Leg Exceptions Scale for Open Weighing (BLES-OW)

Below is a list of client characteristics that therapists sometimes deem important in considering the appropriateness of open weighing (weighing the patient and discussing their actual weight with them).

Please think about when you use CBT for eating disorders. For each characteristic, rate the likelihood that you would or would not openly weigh the patient.

	Very unlikely to weigh openly	Somewhat Unlikely to weigh openly	Somewhat Likely to weigh openly	Very likely to weigh openly
The patient is below the age of 12.				
The patient is between the ages of 12 and 17.				
The patient is older than age 65.				
The patient is an ethnic minority.				
The patient has a co-morbid personality disorder.				
The patient has co-morbid depression.				
The patient has a co-morbid anxiety disorder.				
The patient is currently experiencing significant stressful life events (e.g., divorce, loss of job, etc.).				
The patient is emotionally fragile.				
The patient does not want to be weighed.				
The patient says that being weighed will have negative consequences (e.g., they will restrict or binge).				
The patient has angry outbursts.				
The patient is pregnant.				
The patient has been weighed in therapy before, and did not find it useful.				
The patient has below average intelligence.				
The patient has poor insight into the irrational nature of his or her fear(s).				

Appendix G

Firm Empathy Questionnaire (FEQ)

People see themselves and other people in a range of different ways. Please tell us about yourself by rating how much each of the following statements apply to you. Please tick one box per item.

Number	Item	Never	Rarely	Sometimes	Often	Most or the time	Nearly all the time	Always
1	I am able to put myself in other people's shoes							
2	I am not afraid to assert myself							
3	I feel sad when someone is genuinely upset							
4	I think it is important to make exceptions to the rules in some situations							
5	I find it hard to connect with others on an emotional level							
6	I like to make sure I achieve the goals that I set for myself							
7	I can understand other people's feelings easily							
8	I think it is important to stick to the recommended way of carrying out tasks							
9	I don't like to concern myself with other people's feelings							
10	I believe that it is important to have boundaries							
11	When other people are happy, I share that feeling							
12	I like to make sure others achieve their tasks							
13	Other people’s feelings do not interest me							
14	It matters if people do not do what they were meant to do							
15	I worry about other people who are having a hard time							
16	When rules are violated, there should be consequences							

Appendix H

Core Vignette

You are seeing **[insert name],** a 25 year old patient, for her first CBT treatment session. She has been assessed as having **[insert diagnosis: anorexia nervosa OR bulimia nervosa]**. You've been talking for 15 minutes about her eating habits and have now come to a point in the session where it would be appropriate to weigh her if you were going to. **[Insert level of distress: She appears calm; OR She seems anxious and distressed about things that are going on in her life at the moment; OR She seems anxious and distressed about the prospect of being weighed.]**

Example Vignette:

You are seeing Louise, a 25 year old patient, for her first CBT treatment session. She has been assessed as having anorexia nervosa. You've been talking for 15 minutes about her eating habits and have now come to a point in the session where it would be appropriate to weigh her if you were going to. She seems anxious and distressed about the prospect of being weighed.

Appendix I

Weighing Behaviour Questions

You will now be shown brief descriptions of some patients with eating disorders. CBT clinicians vary in their weighing practices, with some weighing the patient themselves, some arranging for others to do it, and some asking the patient to do it. Similarly, some clinicians weigh their patients 'openly' (showing or telling the patient their weight) and some weigh their patients 'blind' (weighing the patient but not telling them their weight). We are interested in how you would weigh each of the patients in these vignettes when treating them using a cognitive-behavioural therapy approach (CBT).

Please read each description carefully. After each description, answer the questions about what you would do in the session.

[vignette]

On the following scales (0-100% where 0% = never and 100% = all the time), please rate how likely you are to do each of the following regarding weighing [patient name from vignette].

Please only use the 'N/A' option for the final question.

	0%	10%	20%	30%	40%	50%	60%	70%	80%	90%	100%	N/A
Rely on [patient name]'s self-reported weight												
Determine 'by eye' whether [patient name]'s weight causes any concerns												
Weigh [patient name] in the session												
Delay weighing [patient name] until another session												
Ask someone else to weigh [patient name]												
Tell [patient name] her weight, if she has been weighed (If she has not been weighed, please select 'N/A')												

Appendix J

Participant Information and Informed Consent

Participant Information and Informed Consent

Thank you for your interest in our study.

We are investigating what factors influence CBT clinicians' weighing practices with their eating disordered patients. We would like to ask you to take part by reading brief descriptions of some patients with eating disorders, and answering questions about your weighing practices. We also ask you to complete some questionnaires about how you feel and your beliefs about weighing patients.

Your responses are anonymous and confidential and the data will only be kept for the purpose of this research. This research has been approved by the University of Sheffield's Department of Psychology Ethics Committee, and is supervised by Professor Glenn Waller.

If you have any questions or concerns please contact Amy Daglish (adaglish1@sheffield.ac.uk) or Glenn Waller (g.waller@sheffield.ac.uk).

If you have any further concerns, please contact Lindsay Unwin, Secretary to Sheffield University Research Ethics Committee on 0114 22 21443.

I understand that the information that is collected during this study will be confidential

- Yes
- No

I agree to take part in this study

- Yes
- No

In case you decide at a later stage that you do not wish your data to be used, we need you to create a code that will enable us to find your data without asking for your name at this stage. Therefore, please create a code as follows:

Use the first two letters of your mother's maiden name, the day of the month when you were born, and the first two letters of your surname. So if your mother's maiden name was 'Johnson', you were born on the 17th of the month, and your surname is 'Smith', your code will be JO17SM.

Appendix K

Participant Debriefing

Thank you for taking part in this study. If you are interested in knowing more about what the study was investigating, read on...

This study is investigating how patient and clinician characteristics influence CBT clinicians' weighing practices when working with eating disordered patients. We hope that our findings will inform future training to support clinicians to use evidence-based CBT with their eating disordered patients.

Why should I weigh my patient in CBT for eating disorders?

The excerpt below summarizes the reasons for weighing and we hope you find this information useful:

"There are four reasons for CBT therapists to weigh their eating disordered patients - to keep them safe, to understand their eating patterns, to reduce the patient's anxiety and avoidance, and to modify the central cognitive problem at the heart of the eating disorders. In order to address the first two, weighing can be undertaken in many different ways, provided that the clinician is able to monitor the results (e.g., the person doing the weighing communicates that weight to the person delivering the therapy) and as long as any concerns about weight falsification are addressed (e.g., by checking for electrolyte imbalances that might indicate water loading). However, the latter two reasons reflect the cognitive and behavioral combination that is central to CBT for eating disorders, and need to be carried out in an appropriate way. Indeed, it can be concluded that weighing the patient appropriately is necessary for the therapy to be seen as CBT."

From: Waller G., & Mountford, V. (2015) Weighing patients within cognitive-behavioral therapy for eating disorders: how, when and why. *Behavior Research and Therapy. 70,* 1–10. doi: 10.1016/j.brat.2015.04.004.

If you wish to discuss your weighing practices you should contact your clinical supervisor.

If you have any concerns or queries regarding this research, or would like a summary report when the research is complete, please contact Amy Daglish (adaglish1@sheffield.ac.uk) or Glenn Waller (g.waller@sheffield.ac.uk)