Mediators of Transdiagnostic Psychological Treatments for Eating Disorders



Katy A. J. Sivyer, MSc, BA

Jesus College

Thesis submitted for the degree of Doctor of Philosophy at the University of Oxford

Trinity Term 2017

Abstract

Title: Mediators of Transdiagnostic Psychological Treatments for Eating Disorders

Author: Katy A. J. Sivyer,

College: Jesus College

Submitted: Trinity Term 2017

Abstract: Cognitive behaviour therapy and interpersonal psychotherapy are the leading treatments for eating disorders. Little is understood regarding their mechanisms of action. The research described in this thesis investigated the purported mechanisms of action of two transdiagnostic versions of these treatments; enhanced cognitive behaviour therapy (CBT-E) and interpersonal psychotherapy for eating disorders (IPT-ED). A series of mediation studies were embedded within a randomised controlled trial comparing CBT-E and IPT-ED in a transdiagnostic sample of patients with eating disorders. An analytic strategy using multilevel and structural equation modelling was used to assess for statistical mediation. Three of the key purported mediators of action of CBT-E (regularity of eating, frequency of weighing, and frequency of shape checking) and the key purported mediator of IPT-ED (interpersonal problem severity) were studied. Only regularity of eating demonstrated a strong case for being a mediator of the effect of CBT-E (on frequency of binge eating). The findings were inconclusive regarding other putative mediators of the effect of CBT-E, and for interpersonal problem severity being a mediator of the effect of IPT-ED. Limitations of the research included the non-optimal choice of measures and non-optimal timing of measurements for establishing temporal precedence. Future research should investigate the mediating role of both cognitive (e.g. interpretation of weight) and behavioural processes (e.g. frequency of weighing) in CBT-E, and the role of interpersonal functioning in CBT-E and IPT-ED. Research should use daily, or session by session measurement to better assess the temporal relationship between the purported mediator/s and the outcome/s. Experimental designs comparing dismantled versions of treatment would also help determine the relative effects that different treatment procedures have on treatment outcome/s.

Dedication

To those who make my world a better place; my family and my friends. Especially; Stephanie Aplin-Wakefield, Laura Montgomery-Hurrell, Andrew Montgomery, Charlotte Beach, Elizabeth Rooney, Michael Rooney, Stephen Edwards, and Abel Toledano Infantes. Thank you for your patience over the last few months and for your love and support always.

Acknowledgements

I would like to thank my supervisors, Professor Christopher Fairburn and Dr Elizabeth Allen, without whom I would not have been able to do this research. Your advice and support has been invaluable.

My thanks also goes to my colleagues at the Centre for Research on Eating Disorders at Oxford (CREDO) for their support throughout my time at CREDO; Dr Suzanne Straebler-Bailey, Dr Rebecca Murphy, Professor Zafra Cooper, Marianne O'Connor, and Emily Rothwell. I would also like to thank the therapists and patients who took part in this research, who I never met but without whom this research would not have been possible.

Additional thanks go to my colleagues from the University of Southampton who have cheered me on the last few months of write-up; Dr Kate Greenwell, Dr Philine Harris and Dr Elena Heber. Also to Richard Elliott, Dr Jessica Bolland, Damien Frost, Angus Stevner, Annina Hessel, Emma Moran, and Olga Batty for their support and encouragement during my DPhil. Finally, to Dr Claire Hart who originally set me on the path of advanced statistics.

Without the love and support of the special people in my life I would not have completed this work. For their support, now and always, I would like to thank; Stephanie Aplin-Wakefield, my amazing auntie, who has been there every step of the way; my adopted family and my best friends, Laura Montgomery-Hurrell, Andrew Montgomery, Charlotte Beach, Luke Middlewick, and Stephen Edwards. Lastly, my love and thanks go to Abel Toledano Infantes, for his apparently infinite patience; thank you for putting up with the ranting, crying, lack of domestic chore completion and for making me laugh when I needed it the most. Also, for making sure there was always an emergency supply of chocolate in the house.

Statement of Authorship

The work described in this thesis was undertaken as part of a programme of research investigating the efficacy of psychological treatments for eating disorders. This research programme is directed by Professor Christopher Fairburn. This thesis examines data collected from a randomised controlled trial comparing two psychological treatments for eating disorders. This trial was designed so that a series of longitudinal mediation studies were embedded in the main treatment efficacy trial (Fairburn et al., 2015; Murphy, Cooper, Hollon, & Fairburn, 2009).

The specific contribution of the candidate to this research was as follows:

Data Cleaning

I was responsible for checking the data entry for the mediation studies (including undertaking a 20% data quality spot check of the data). Where data had not yet been entered, I carried out the data entry myself.

Planning of Statistical Analyses

I planned and developed the analytic strategy used in this thesis.

Data Analysis

I undertook all the analyses described in this thesis.

The Thesis

The thesis has been prepared and written entirely by myself.

Presentations

I was responsible for the following presentations resulting from this work:

Sivyer, K.A.J., Murphy, R., Allen, E. Cooper, Z & Fairburn, C. G. (accepted, 2017, July). Assessing Regularity of Eating as a Purported Mediator of the Effect of Enhanced Cognitive Behaviour Therapy on Frequency of Binge Eating. Oral presentation at the British Association for Behavioural and Cognitive Psychotherapies (BABCP) Conference, Manchester, UK.

Sivyer, K.A.J., Murphy, R., Allen, E. Cooper, Z & Fairburn, C. G. (accepted, 2017, July). Frequency of Shape Checking as a Purported Mediator of the Effect of Enhanced Cognitive Behaviour Therapy on Concern about Shape in Eating Disorders. Poster presentation at the British Association for Behavioural and Cognitive Psychotherapies (BABCP) Conference, Manchester, UK.

Sivyer, K.A.J., Murphy, R., Allen, E. Cooper, Z & Fairburn, C. G. (2016, September). Complexities of running mediation analysis in a randomised controlled trial with two active treatment arms. Oral presentation at the Royal Statistical Society International Conference

Sivyer, K.A.J., Murphy, R., Allen, E. Cooper, Z & Fairburn, C. G. (2016, June). Regular Eating as a Mediator of Enhanced Cognitive Behaviour for Eating Disorders. Poster presentation at the British Association for Behavioural and Cognitive Psychotherapies (BABCP) Conference, Belfast, UK. [Awarded a BABCP Excellence Award for Best Poster]

Contents

Abstract iii
Dedication iv
Acknowledgements v
Statement of Authorship vi
Contents viii
List of Tables xiii
List of Figures xv
List of Appendices xix
List of Abbreviations xxii
Introduction 1
Chapter 1 4
Chapter 1. Development of Enhanced Cognitive Behaviour Therapy and Interpersonal Psychotherapy for Eating Disorders 4
1.1. Introduction 4
1.2. Cognitive Therapy and Interpersonal Psychotherapy for Depression 4
1.2.1. Description of Cognitive Therapy for Depression 5
1.2.2. Description of Interpersonal Psychotherapy for Depression 7
1.2.3. Current Evidence-Base for Cognitive Therapy and Interpersonal Psychotherapy for Depression 9
1.3. Cognitive Behaviour Therapy and Interpersonal Psychotherapy for Bulimia Nervosa 11
1.3.1. Description of Cognitive Behaviour Therapy for Bulimia Nervosa 12
1.3.2. Description of Interpersonal Psychotherapy for Bulimia Nervosa 15
1.3.3. Current Evidence-Base for Cognitive Behaviour Therapy and Interpersonal Psychotherapy for Bulimia Nervosa 17
1.4. Enhanced Cognitive Behaviour Therapy and Interpersonal Psychotherapy for Eating Disorders 18
1.4.1. Description of Enhanced Cognitive Behaviour Therapy 19
1.4.2. Description of Interpersonal Psychotherapy for Eating Disorders 24
1.4.3. Current Evidence-Base for Enhanced Cognitive Behaviour Therapy and Interpersonal Psychotherapy for Eating Disorders 27
1.5. Conclusion 39
Chapter 2 41
Chapter 2. Research into the Mechanisms of Action of Enhanced Cognitive Behaviour Therapy and Interpersonal Psychotherapy for Eating Disorders 41
2.1. Introduction 41
2.1.1. Definition of Terms 42

2.1.2. Research Designs Used to Identify Mechanisms of Treatment Action 45
2.2. Non-Specific Predictors of Treatment Outcome in CBT and IPT in Eating Disorders 47
2.3. Moderators of Treatment Outcome in CBT and IPT in Eating Disorders 49
2.4. Mechanisms of Action of CBT and IPT in Eating Disorders 49
2.4.1. Observational Studies 50
2.4.2. Experimental Studies 51
2.5. Conclusion 62
Chapter 3 64
Chapter 3. Research Design 64
3.1. Introduction 64
3.2. Overview of Main Efficacy Trial 64
3.2.1. Research Design and Procedure 64
3.2.2. Power 68
3.2.3. Statistical Analysis and Results of Main Trial 69
3.3. Mediation Studies 69
3.3.1. Research Design and Procedure 69
3.4. Aims of Thesis 75
Chapter 4 77
Chapter 4. Statistical Methods for Investigating Purported Mechanisms of Action of Treatment 77
4.1. Traditional Approaches to Statistical Mediation Analysis 77
4.1.1. Causal Steps Approach 77
4.1.2. Structural Equation Modelling 91
4.1.3. Statistical Power to Detect a Mediated Effects 108
4.2. Modern Approaches to Statistical Mediation Analysis 109
4.3. Analytic Strategy 110
4.3.1. Data Preparation 112
4.3.2. Study A: A Preliminary Investigation using Multilevel Modelling for Statistical Mediation 114
4.3.3. Study B: An In-Depth Analysis using Autoregressive Structural Equation Modelling for Statistical Mediation 118
4.3.4. Study C: Testing an Alternative Model using Autoregressive Structural Equation Modelling for Statistical Mediation 120
4.3.5. Study D: An Investigation into the Relationship between Change in the Mediator and Change in Outcome using Latent Growth Curve Modelling 120

4.3.6. Criteria for Accepting a Purported Mediator as a Potential Mechanism of Action of Treatment 121
Chapter 5 124
Chapter 5. Study 1: Regularity of Eating as a Mediator of the Effect of Enhanced Cognitive Behaviour Therapy on Frequency of Binge Eating 124
5.1. Background 124
5.2. Aim/Hypothesis 125
5.3. Method 126
5.3.1. Experimental Procedure 126
5.3.2. Study Measures 127
5.3.3. Analytic considerations 128
5.4. Results 129
5.4.1. Non-Normality and Missing Data 129
5.4.2. Study 1a: A Preliminary Investigation using Multilevel Modelling for Statistical Mediation 129
5.4.3. Study 1b: An In-Depth Analysis using Autoregressive Structural Equation Modelling for Statistical Mediation 134
5.4.4. Study 1c: Testing an Alternative Model: Frequency of Binge Eating (the Outcome) as Potential Mediator of Change in Regularity of Eating (the Purported Mediator) 139
5.4.5. Study 1d: An Investigation into the Relationship between Change in the Mediator and Change in Outcome using Latent Growth Curve Modelling 144
5.5. Summary of Results 150
Chapter 6 151
Chapter 6. Study 2: Frequency of Weighing as a Mediator of the Effect of Enhanced Cognitive Behavioural Therapy on Concern about Weight 151
6.1. Background 151
6.2. Aim / Hypothesis 152
6.3. Method 153
6.3.1. Experimental Procedure 153
6.3.2. Study Measures 153
6.3.3. Analytic considerations 154
6.4. Results 155
6.4.1. Non-Normality and Missing Data 155
6.4.2. Study 2a: A Preliminary Investigation using Multilevel Modelling for Statistical Mediation 156
6.4.3. Study 2b: An In-Depth Analysis using Autoregressive Structural Equation Modelling for Statistical Mediation 161

6.5. Summary of Results 166
Chapter 7 167
Chapter 7. Study 3: Frequency of Shape Checking as a Mediator of the Effect of Enhanced Cognitive Behaviour Therapy on Concern about Shape 167
7.1. Background 167
7.2. Aim/Hypothesis 168
7.3. Method 169
7.3.1. Experimental Procedure 169
7.3.2. Study Measures 169
7.3.3. Analytic considerations 170
7.4. Results 175
7.4.1. Non-Normality and Missing Data 175
7.4.2. Implementation of the Shape checking intervention 175
7.4.3. Study 3a: A Preliminary Investigation using Multilevel Modelling for Statistical Mediation 176
7.4.4. Study 3b: An In-Depth Analysis using Autoregressive Structural Equation Modelling for Statistical Mediation 181
7.5. Summary of Results 186
Chapter 8 187
Chapter 8. Study 4: Interpersonal Problem Severity as a Mediator of the Effect of Interpersonal Psychotherapy for Eating Disorders on Eating Disorder Psychopathology 187
8.1. Background 187
8.2. Aim/Hypothesis 187
8.3. Method 189
8.3.1. Experimental Procedure 189
8.3.2. Study Measures 190
8.3.3. Analytic considerations 191
8.4. Results 192
8.4.1. Non-Normality and Missing Data 192
8.4.2. Study 4a: A Preliminary Investigation using Multilevel Modelling for Statistical Mediation 192
8.4.3. Study 4b: An In-Depth Analysis using Autoregressive Structural Equation Modelling for Statistical Mediation 197
8.5. Summary of Results 202
Chapter 9 203
Chapter 9. Discussion 203

9.1. Overview 203
9.2. Key Findings and Implications 204
9.2.1. Purported Mediators of Enhanced Cognitive Behaviour Therapy 204
9.2.2. Purported Mediators of Interpersonal Psychotherapy for Eating Disorders 215
9.3. Methodological Considerations and Implications 219
9.3.1. Research Design 219
9.3.2. Statistical Analysis 222
9.4. General Discussion 229
9.5. Conclusion 231
References 233
Appendices 243

List of Tables

Table 1: Summary of interpersonal problem areas and their treatment strategies..... 10
Table 2: Overview of research investigating the effectiveness of enhanced cognitive behaviour therapy (CBT-E) and interpersonal psychotherapy for eating disorders (IPT-ED)..... 32
Table 3: Summary of remission rates of psychopathology (score <1SD above community norm for eating disorder psychopathology) reported in studies in Table 2 37
Table 4: Summary of remission rates of behavioural symptomatology (cessation of bingeing and purging, body mass index ≥ 18.5) reported in studies in Table 2. 37
Table 5: Non-specific predictors of treatment outcome in eating disorders identified in the Vall and Wade (2015) meta-analysis..... 48
Table 6: Overview of treatment research investigating treatment-specific purported mechanisms of cognitive behaviour therapy (CBT) and interpersonal psychotherapy (IPT) in eating disorders..... 54
Table 7: Treatment timings of enhanced cognitive behaviour therapy (CBT-E) and interpersonal psychotherapy for eating disorders (IPT-ED)..... 71
Table 8: Timing of Measures 72
Table 9: Summary of Four Common Fit Indices, their Cut-Offs and their Relative Strengths and Limitations..... 97
Table 10: Breakdown of number of items in dataset by measure and number of repeated measurements 113
Table 11: Comparison of the cumulative indirect and total effects of the intervention (CBT-E) on frequency of binge eating (BE) each week controlling for shared measurement variance, depression, eating disorder diagnosis, presence of binge eating in the month prior to starting treatment, age and weight concern at baseline 138
Table 12: Comparison of the cumulative indirect and total effects of the intervention (CBT-E) on regularity of eating (RE) each week controlling for shared measurement variance, depression, eating disorder diagnosis, presence of binge eating in the month prior to starting treatment, age and weight concern at baseline 143
Table 13: Comparison of the indirect (via change in regularity of eating only) and total effects of the intervention (CBT-E) on frequency of binge eating (BE) during the first four weeks of treatment, controlling for depression, eating disorder diagnosis, presence of binge eating in the month prior to starting treatment, age and weight concern at baseline..... 149
Table 14: Comparison of the cumulative indirect and total effects of the intervention (CBT-E) on concern about weight each week in non-frequent weighers, controlling for shared measurement variance, self-esteem and weight concern at baseline 165
Table 15: Examples of coding schemes for personalisation of treatment variable..... 172
Table 16: Comparison of the cumulative indirect and total effects of the shape checking intervention on concern about shape each week controlling for shared

measurement variance, treatment allocation, eating disorder duration, self-esteem, age, weight concern, depression, importance of shape, frequency of checking and concern about shape at baseline 185
Table 17: Comparison of the cumulative indirect and total effects of IPT-ED on interpersonal problem severity at each study time point controlling for shared measurement variance, eating disorder duration, depression and self-esteem at baseline..... 201

List of Figures

Figure 1: The cognitive behavioural model of bulimia nervosa (modified from Fairburn, Marcus, & Wilson, 1993). 14
Figure 2: The enhanced cognitive behavioural model of eating disorders (modified from Fairburn (2008)) 22
Figure 3: The interpersonal model of eating disorders 26
Figure 4: Comparison of remission rates in CBT-E and IPT-ED in the CREDO efficacy trials (reproduced with permission from (Fairburn et al., 2015)) 38
Figure 5: Illustration of the impact of a non-specific prediction on treatment outcome 42
Figure 6: Illustration of the impact of a moderator on treatment outcome 43
Figure 7: Illustration of the relationships between treatment, mediator and outcome 43
Figure 8: Regression model 1 in the causal steps approach 78
Figure 9: Regression model 2 in the causal steps approach 78
Figure 10: Regression model 3 in the causal steps approach 79
Figure 11: Regression model for testing treatment moderators 80
Figure 12: Multilevel Modelling Equations for a 2-1-1 mediation model in which treatment is randomised and both the purported mediator and the outcome are repeated measures. 82
Figure 13: Example Data Structure of a 2-1-1 Multilevel Mediation Model 83
Figure 14: Illustration of how treatment's effect on the purported mediator and consequently the outcome might vary between treatments 86
Figure 15: Modified regression model 3 in updated causal steps approach testing for an interactive effect between treatment and the purported mediator 87
Figure 16: Decomposition of the coefficients (represented as paths) needed to calculate the total, direct and indirect effects of treatment 88
Figure 17: Path analysis models for full and partial mediation 92
Figure 18: Example of an autoregressive model of Outcome O at time points 1-3 98
Figure 19: Different specifications of three panel autoregressive models for full and partial mediation when treatment allocation has been randomised at baseline (Models A-C, and D-F, respectively). Models A and C illustrate longitudinal mediation controlling for covariance between the purported mediator and the outcome within the same time point. Models B and D depict longitudinal and contemporaneous relationships between the purported mediator and the outcome. Models C and E demonstrate cross-lagged relationships between both the purported mediator and the outcome. 101
Figure 20: Example latent growth curve model for outcome O over three time points (error terms omitted) 103
Figure 21: Parallel Process Latent Growth Curve Models for three time points (error terms omitted) 107
Figure 22: Illustration of the purported mediational relationship between enhanced cognitive behaviour therapy (CBT-E), regularity of eating and frequency of binge eating 125

Figure 23: Line graph plotting the observed means (smoothed curve) of regularity of eating and frequency of binge eating in IPT-ED 131
Figure 24: Line graph plotting the observed means (smoothed curve) of regularity of eating and frequency of binge eating in CBT-E 131
Figure 25: Coefficient plots with 95% bias-corrected confidence intervals (1000 bootstrap resamples) comparing the effect of the intervention (CBT-E) on regularity of eating and the frequency of binge eating when regularity of eating was / was not included as a mediator. Impact of treatment allocation represents an interaction between treatment allocation and treatment week. Controlling for the effects of eating disorder diagnosis, depression, weight concern, presence of binge eating in the month prior to starting treatment, and age at baseline. 132
Figure 26: Autoregressive structural equation model of the impact of treatment on frequency of binge eating (BE) via regularity of eating (RE). Numeral indicates treatment week. Controlling for shared measurement variance, depression, eating disorder diagnosis, presence of binge eating in the month prior to starting treatment, age and weight concern at baseline (standardised estimates in brackets). Regularity of eating denoted as 'RE' and frequency of binge eating denoted as 'BE', where numeral indicates treatment week. Significance denoted by * based on bootstrapped bias-corrected confidence intervals (1000 resamples). Missing data accounted for using full information maximum likelihood estimation. 137
Figure 27: Alternative autoregressive structural equation model of the impact of treatment on regularity of eating (RE) as an outcome via frequency of binge eating (BE) as an alternative mediator. Numeral denotes treatment week. Controlling for shared measurement variance, depression, eating disorder diagnosis, presence of binge eating in the month prior to starting treatment, age and weight concern at baseline (standardised estimates in brackets). Significance denoted by * based on bootstrapped bias-corrected confidence intervals (1000 resamples). Missing data accounted for using full information maximum likelihood estimation. 142
Figure 28: Parallel process latent growth curve model examining the impact of treatment allocation on frequency of binge eating (BE) via regularity of eating (RE). Numeral indicates treatment week. Controlling for depression, eating disorder diagnosis, presence of binge eating in the month prior to starting treatment, age and weight concern at baseline (standardised estimates in brackets). Significance denoted by * based on bias-corrected confidence intervals (1000 bootstrap resamples). 148
Figure 29: Illustration of the purported mediational relationship between enhanced cognitive behaviour therapy (CBT-E), frequency of weighing and concern about weight 152
Figure 30: Line graph plotting the observed means (smoothed curve) of frequency of weighing and concern about weight in non-frequent weighers in the IPT-ED and CBT-E 158
Figure 31: Line graph plotting the observed means (smoothed curve) of frequency of weighing and concern about weight in frequent weighers in the IPT-ED and CBT-E 158

Figure 32: Coefficient plots with 95% bias-corrected confidence intervals (1000 bootstrap resamples) comparing the effect of CBT-E on frequency of weighing and concern about weight when frequency of weighing was / was not included as a mediator. Impact of treatment allocation represents an interaction between treatment allocation and treatment week. Controlling for the effects of baseline frequent weigher status, self-esteem and weight concern at baseline..... 159
Figure 33: Autoregressive structural equation model of the impact of treatment on concern about weight (CS) via frequency of weighing (FW) in frequent and non-frequent weighers. Numeral indicates treatment week. Controlling for shared measurement variance, self-esteem and weight concern at baseline (standardised estimates in brackets). Effect of treatment allocation and baseline frequent weigher status on frequency of weighing modelled as an interaction. Significance denoted by * based on bootstrapped bias-corrected confidence intervals (1000 resamples). Missing data accounted for using full information maximum likelihood estimation..... 164
Figure 34: Illustration of the purported mediational relationship between enhanced cognitive behaviour therapy (CBT-E), frequency of shape checking and concern about shape 168
Figure 35: Illustration of implementation of shape checking intervention and the impact this has on randomisation 174
Figure 36: Line graph plotting the observed means (smoothed curve) of frequency of shape checking and concern about shape in the control group in weeks 4 to 20 of treatment 178
Figure 37: Line graph plotting the observed means (smoothed curve) of frequency of shape checking and concern about shape in the intervention group from 4 weeks before the shape checking intervention was implemented until 10 weeks after 178
Figure 38: Coefficient plots with 95% bias-corrected confidence intervals (1000 bootstrap resamples) comparing the effect of the shape checking intervention on frequency of shape checking and concern about shape when frequency of shape checking was / was not included as a mediator. Impact of treatment allocation represents an interaction between treatment allocation and treatment week. Controlling for the effects of eating disorder duration, age, weight concern, self-esteem, depression and importance of shape at baseline. 179
Figure 39: Autoregressive structural equation model of the impact of the shape checking intervention on concern about shape (CS) via frequency of shape checking, (FC). Numeral indicates intervention week. Controlling for shared measurement variance, treatment allocation, eating disorder duration, age, weight concern, self-esteem, depression, importance of shape, frequency of checking and worry about shape at baseline (standardised estimates in brackets). Significance denoted by * based on bootstrapped bias-corrected confidence intervals (1000 resamples). Missing data accounted for using full information maximum likelihood estimation..... 184

Figure 40: Illustration of the purported mediational relationship between interpersonal psychotherapy for eating disorders (IPT-ED), interpersonal problem severity and eating disorder psychopathology 188
Figure 41: Line graph plotting the observed means (smoothed curve) of interpersonal problem severity and eating disorder psychopathology in CBT-E during treatment and follow-up 194
Figure 42: Line graph plotting the observed means (smoothed curve) of interpersonal problem severity and eating disorder psychopathology in IPT-ED during treatment and follow-up 194
Figure 43: Coefficient plots with 95% bias-corrected confidence intervals (1000 bootstrap resamples) comparing the effect of IPT-ED on interpersonal problem severity and eating disorder psychopathology when interpersonal problem (IP) severity was / was not included as a mediator. Impact of treatment allocation represents an interaction between treatment allocation and treatment week. Controlling for the effects of eating disorder diagnosis, eating disorder duration, self-esteem and depression at baseline. 195
Figure 44: Autoregressive structural equation model of the impact of treatment on eating disorder psychopathology (EDP) via interpersonal problem severity (IPS). Numeral indicates study time point (0=baseline, 1=end of treatment, 2=20 week follow-up, 3=40 week follow-up). Controlling for shared measurement variance, eating disorder diagnosis, eating disorder duration, self-esteem, and depression at baseline (standardised estimates in brackets). Significance denoted by * based on bootstrapped bias-corrected confidence intervals (1000 resamples). Missing data accounted for using full information maximum likelihood estimation. 200

List of Appendices

Appendix 1: Main Efficacy Trial Paper (Fairburn et al., 2015) 243
Appendix 2: Mediation Trial Pre-Paper (Murphy et al., 2009) 250
Appendix 3: Regular Eating Rating Scale (RERS) 255
Appendix 4: Regular Eating Rating Scale (RERS) and Binge Eating Rating Scale (BERS) Rating Form 256
Appendix 5: Binge Eating Rating Scale 257
Appendix 6: Eating Questionnaire (EQ) 258
Appendix 7: Interpersonal Questionnaire (IQ) 261
Appendix 8: Interpersonal Interview (IPI) 262
Appendix 9: Interpersonal Problem Form 270
Appendix 10: Intervention Record Sheet for Enhanced Cognitive Behaviour Therapy 272
Appendix 11: Intervention Record Sheet for Interpersonal Psychotherapy for Eating Disorders 273
Appendix 12: Monitoring Record for Enhanced Cognitive Behaviour Therapy 274
Appendix 13: Monitoring Record Interpersonal Psychotherapy for Eating Disorders 275
Appendix 14: Stata code for study 1a multilevel models and their coefficient plots .. 276
Appendix 15: Output for study 1a, step 1. Multi-level model of the impact of the intervention (CBT-E) on the purported mediator (regularity of eating) 279
Appendix 16: Output for study 1a, step 2. Multi-level model of the impact of the intervention (CBT-E) on the outcome (frequency of binge eating) 280
Appendix 17: Output for study 1a, step 3. Multi-level model of impact of the purported mediator (regularity of eating) on the outcome (frequency of binge eating) 281
Appendix 18: Output for study 1a, step 4. Multi-level model of impact of the intervention (CBT-E) on the outcome (frequency of binge eating) when accounting for the purported mediator (regularity of eating) 282
Appendix 19: Stata code for study 1b autoregressive structural equation model 283
Appendix 20: Output for study 1b. Autoregressive structural equation model of the relationships between the intervention (CBT-E), the purported mediator (regularity of eating; RE) and the outcome (frequency of binge eating; BE) ... 285
Appendix 21: Stata code for study 1c alternative autoregressive structural equation model 287
Appendix 22: Output for study 1c. Alternative Autoregressive structural equation model of the relationships between the intervention (CBT-E), frequency of binge eating as an alternative mediator (BE) and regularity of eating as an alternative outcome (RE) 289
Appendix 23: Stata code for study 1d latent growth curve model 291
Appendix 24: Output for study 1d. Parallel process latent growth curve model of the relationships between the intervention (CBT-E), change in the purported mediator (regularity of eating; RE) and change in the outcome (frequency of binge eating; BE) 293
Appendix 25: Stata code for study 2a multilevel models and their coefficient plots .. 295

Appendix 26: Output for study 2a, step 1. Multi-level model of the impact of the intervention (CBT-E) on the purported mediator (frequency of weighing) 298
Appendix 27: Output for study 2a, step 2. Multi-level model of the impact of the intervention (CBT-E) on the outcome (concern about weight) It would be good to see the estimates of the SD of the random effects and of the residual errors 299
Appendix 28: Output for study 2a, step 3. Multi-level model of impact of the purported mediator (frequency of weighing) on the outcome (concern about weight) 300
Appendix 29: Output for study 2a, step 4. Multi-level model of impact of the intervention (CBT-E) on the outcome (concern about weight) when accounting for the purported mediator (frequency of weighing) 301
Appendix 30: Stata code for study 2b autoregressive structural equation model 302
Appendix 31: Output for study 2b. Autoregressive structural equation model of the relationships between the intervention (CBT-E), the purported mediator (frequency of weighing; FW) and the outcome (concern about weight; CW). 305
Appendix 32: Stata code for study 3a multilevel models and their coefficient plots .. 309
Appendix 33: Output for study 3a, step 1. Multi-level model of the impact of the shape checking intervention on the purported mediator (frequency of shape checking) 312
Appendix 34: Output for study 3a, step 2. Multi-level model of the impact of the shape checking intervention on the outcome (concern about shape)..... 313
Appendix 35: Output for study 3a, step 3. Multi-level model of impact of the purported mediator (frequency of shape checking) at time (T) on the outcome (concern about shape) at time (T) 314
Appendix 36: Output for study 3a, step 4. Multi-level model of impact of the shape checking intervention on the outcome (concern about shape) when accounting for the purported mediator (frequency of shape checking) within the same treatment week 315
Appendix 37: Stata code for study 3b autoregressive structural equation model 316
Appendix 38: Output for study 3b. Autoregressive partial mediation model of the relationships between the shape checking intervention, frequency of shape checking at the previous week and within the same week, and concern about shape at time (T) using full information maximum likelihood estimation for missing data 319
Appendix 39: Stata code for study 4a multilevel models and their coefficient plots .. 324
Appendix 40: Output for study 4a, step 1. Multi-level model of the impact of IPT-ED on the purported mediator (interpersonal problem severity)..... 327
Appendix 41: Output for study 4a, step 2. Multi-level model of the impact of IPT-ED on the outcome (eating disorder psychopathology)..... 328
Appendix 42: Output for study 4a, step 3. Multi-level model of the impact of the purported mediator (interpersonal problem severity) on the outcome (eating disorder psychopathology) 329
Appendix 43: Output for study 4a, step 4. Multi-level model of the impact of IPT-ED on the outcome (eating disorder psychopathology) when accounting for the purported mediator (interpersonal problem severity) within the same time point..... 330

Appendix 44: Stata code for study 4b autoregressive structural equation model 331
Appendix 45: Output for study 4b. Autoregressive partial mediation model of the relationships IPT-ED, interpersonal problem severity at the previous time point and within the same time point, and eating disorder psychopathology using full information maximum likelihood estimation for missing data 333

List of Abbreviations

AIC = Akaike Information Criterion

AN = Anorexia Nervosa

BDI-II = Beck Depression Inventory II

BED = Binge Eating Disorder

BERS = Binge Eating Rating Scale

BMI = Body Mass Index

BN = Bulimia Nervosa

BPD = Borderline Personality Disorder

CBT = Cognitive Behaviour Therapy

CBT-BN = Cognitive Behaviour Therapy for Bulimia Nervosa

CBT-E = Enhanced Cognitive Behaviour Therapy

CBT-Eb = Enhanced Cognitive Behaviour Therapy (broad version)

CBT-Ef = Enhanced Cognitive Behaviour Therapy (focused version)

CFI = Comparative Fit Index

Chi2 = Chi-Square

CREDO = Centre for Research on Eating Disorders at Oxford

CT = Cognitive Therapy

ED = Eating Disorder

EDE = Eating Disorder Examination

EDNOS = Eating Disorder Not Otherwise Specified

EME = Efficacy and Mechanism Evaluation

EQ = Eating Questionnaire

FBT = Family=Based Therapy

ICAT = Integrative Cognitive-Affective Therapy

IPI = Interpersonal Interview

IPT = Interpersonal Psychotherapy

IPT-BN = Interpersonal Psychotherapy for Bulimia Nervosa

IPT-ED = Interpersonal Psychotherapy for Eating Disorders

IQ = Interpersonal Questionnaire

IRS = Intervention Record Sheet

MLM = Multilevel Modelling

RCT = Randomised Controlled Trial

RERS = Regular Eating Rating Scale

RMSEA = Root Mean Square Error of Approximation

RSES = Rosenberg Self-Esteem Scale

SEM = Structural Equation Modelling

Introduction

This research investigates potential mechanisms of action of two psychological treatments for eating disorders; enhanced cognitive behavioural therapy (CBT-E) and interpersonal psychotherapy for eating disorders (IPT-ED). The thesis addresses two main questions. Firstly, it addresses the clinical question regarding whether the hypothesised mechanisms of action of CBT-E and IPT-ED are observable during treatment, if they relate meaningfully to outcome and whether they potentially mediate the effect of treatment on the outcome. Secondly, it addresses the statistical question regarding potential analytic strategies for addressing the challenges of undertaking statistical mediation in the context of a randomised controlled trial of two active treatments and treatment personalisation. The thesis is split into four main sections. Section 1 focuses on the background to the thesis. Section 2 describes the study methods and overall analytic strategy. Section 3 discusses the studies undertaken as part of the thesis. Finally, section 4 discusses the results of the research and its implications.

Overview of Chapters

Section 1 – Background

Chapter 1. Development of Enhanced Cognitive Behaviour Therapy and Interpersonal Psychotherapy for Eating Disorders

This chapter investigates the origins and development of CBT-E and IPT-ED, their theoretical underpinnings, the nature of the treatments, and their current

evidence-base.

Chapter 2. Research into the Mechanisms of Action of Enhanced Cognitive Behaviour Therapy and Interpersonal Psychotherapy for Eating Disorders

This chapter defines relevant terms in this area (non-specific predictor, moderator, mediator and mechanism), and explores current research into the non-specific predictors, moderators and mediators/mechanisms of CBT-E and IPT-ED.

Section 2: Study Methods

Chapter 3. Research Design

This chapter describes the overall research design.

Chapter 4. Statistical Methods for Investigating Purported Mechanisms of Action of Treatment

This chapter examines different methods for undertaking statistical mediation analysis, and their relative strengths and weaknesses. It ends with a description of the analytic strategy taken in this thesis.

Section 3: Results

Chapter 5. Regularity of Eating as a Mediator of the Effect of CBT-E on Frequency of Binge Eating

This chapter investigates regularity of eating as a purported mediator of the effect of CBT-E on frequency of binge eating compared to IPT-ED.

Chapter 6. Frequency of Weighing as a Mediator of the Effect of CBT-E on Concern about Weight

This chapter investigates frequency of weighing as a purported mediator of the effect of CBT-E on concern about weight compared to IPT-ED.

Chapter 7. Frequency of Shape Checking as a Mediator of the Effect of CBT-E on Concern about Shape

This chapter investigates frequency of shape checking as a purported mediator of the effect of CBT-E on concern about shape compared to IPT-ED. It also incorporates methods for dealing with treatment personalisation in statistical mediation analyses.

Chapter 8. Interpersonal Problem Severity as a Mediator of the Effect of IPT-ED on Eating Disorder Psychopathology

This chapter investigates interpersonal problem severity as the key purported mediator of the effect of IPT-ED on eating disorder psychopathology compared to CBT-E.

Section 4: Discussion

Chapter 9. Discussion

This chapter summarises and discusses the results from the three studies on purported mediators of CBT-E and, and the one study into the key purported mediator of IPT-ED. This chapter has two main purposes; a) a discussion of the clinical implications of the research; and b) a critique of the methods used and their limitations. Areas of future research are also discussed.

Chapter 1. Development of Enhanced Cognitive Behaviour Therapy and Interpersonal Psychotherapy for Eating Disorders

1.1. Introduction

The origins of enhanced cognitive behaviour therapy (CBT-E) and interpersonal psychotherapy for eating disorders (IPT-ED) date back to the 1960s/1970s during which time two key treatments were being developed. These were cognitive therapy (CT), later also known as cognitive behaviour therapy (CBT) (Beck, 1970; Beck, Rush, Shaw, & Emery, 1984, 1979), and interpersonal psychotherapy (IPT) (Klerman, Weissman, Rounsaville, & Chevron, 1984), both of which were originally designed as treatments for depression.

1.2. Cognitive Therapy and Interpersonal Psychotherapy for Depression

In the 1960s psychoanalytic therapy was the dominant treatment for depression in the USA. However, its evidence-base was poor (Beck et al., 1984, 1979; Weissman, 2006). Whilst the effects of antidepressant medication looked promising, it was not clear what contribution psychotherapy made to treatment outcome, and there was concern that it might undermine the positive effects of medication (Weissman, 2006). CT and IPT developed in the context of these short-comings; CT due to profound changes in the theoretical conceptualisation of depression, and IPT from a more practical basis;

namely, the desire to run a randomised controlled trial (RCT) of antidepressant medication against psychotherapy (Beck, 1970; Weissman, 2006). Whilst IPT was Initially developed as a brief, well-defined supportive psychotherapy, its effectiveness led to its more thorough refinement culminating in the publication of a detailed treatment manual in 1984 (Klerman et al., 1984). Both treatments had some features in common. First, they both focused on resolving current difficulties, rather than exploring the unconscious origins of patients' psychological problems (Beck, 1970; Weissman, 2007). Second, both were well-specified, short-term and time-limited (Beck et al., 1984, 1979; Klerman et al., 1984). However, the two treatments differed in theoretical orientation, evolving from distinct lines of empirical research (Beck et al., 1984; Weissman, 2006).

1.2.1. Description of Cognitive Therapy for Depression

1.2.1.1. Theoretical Model

CT derived from Beck's (1963) cognitive model of depression, which emphasised the role of cognitive structures and processes in this disorder (Beck, 1970). Beck (1963) identified that patients with depression demonstrated faulty information processing (e.g. over-generalisation, black and white thinking) and were inclined towards negative interpretations regarding themselves, the world around them and their future. These interpretations were often experienced as negative 'automatic' thoughts that the patient believed to be accurate representations of reality (Beck, 1963). Beck (1963) hypothesised that these biases underpinned the patients' depression, with negative interpretations triggering negative affect, reinforcing negative conclusions regarding themselves, the world and the future, and resulting in a negative spiral of depression.

He suggested that these biases were driven by underlying beliefs (cognitive structures called 'schema'), which acted as templates for interpreting events based on past experiences, and became activated when triggered by relevant cues in the person or their environment (Beck, 1963). He proposed that correcting patients' negative thoughts and beliefs might ameliorate their depression (Beck, 1964, 1970).

1.2.1.2. The Treatment

The treatment that was derived from this model typically consisted of about 20 sessions of 50 minutes each over approximately 15 weeks, with patients undertaking planned activities ('homework') between therapy sessions (Beck et al., 1984, 1979). Sessions occurred twice-weekly for the first four weeks of treatment for moderate-severely depressed patients, and then once a week thereafter, until the end of treatment, when it was recommended to leave longer gaps between sessions to allow a "tapering off" of therapy (Beck et al., 1984, 1979).

CT had three informal stages of treatment. The first consisted of an orientation and activation phase. This focused on explaining the nature and rationale of the treatment and providing psychoeducation about the disorder. It also had a behavioural focus, predominantly concentrating on increasing patients' levels of daily activity (known as 'behavioural activation') in order to address some of the behavioural and motivational difficulties associated with depression (e.g. withdrawal, immobility) before moving onto more cognitive work. Patients' significant others were invited to attend an interview at this stage to gain additional information about the patients' depression. The main body of treatment (stage two) entailed helping patients identify the cognitive biases that were maintaining their depression, and helping them to assess the

legitimacy of their negative thoughts and beliefs using rationalisation and behavioural experiments designed to test their suppositions (known as 'cognitive restructuring') (Beck et al., 1984, 1979). Patients were asked to record their thoughts, feelings and behaviours and the context in which they occurred ('self-monitoring') from early on in treatment to aid identification of thoughts and beliefs contributing to their depression (Beck et al., 1984, 1979). The final stage of treatment addressed relapse prevention by examining how the patient's cognitive biases and assumptions made the patient vulnerable to depression (Beck et al., 1984, 1979).

1.2.2. Description of Interpersonal Psychotherapy for Depression

1.2.2.1. Theoretical Model

Whilst CT emphasised cognitive and behavioural processes, IPT focussed on the interpersonal context of the depression (Weissman, 2006). In particular, adverse life circumstances and events, such as changes in interpersonal roles (e.g., divorce, changing job) and poor quality of close interpersonal relationships were viewed as being implicated in triggering and maintaining depressive episodes and there was evidence that this was the case (Paykel et al., 1969). Conversely, social support appeared to be protective against developing such problems (Brown & Harris, 1978). Klerman and colleagues (1984) hypothesised that interpersonal problems could act as antecedents to depression by triggering the onset or relapse of depressive episodes, or be consequences of the illness through depression's effect on patients' ability to function in the interpersonal context, creating new interpersonal problems or exacerbating existing ones. These interpersonal problems were then seen to maintain the depression. On this basis Klerman and colleagues (1984) predicted that if poor

interpersonal relationships could trigger or exacerbate depression, then targeting interpersonal relationships in treatment might improve patients' depression.

1.2.2.2. The Treatment

The treatment that was derived from this perspective consisted of approximately 16-20 weekly sessions over 20 weeks (Klerman et al., 1984). Each session lasted approximately an hour and was largely unstructured, with session content increasingly being patient-led, rather than therapist-led, as treatment progressed (Klerman et al., 1984). There was no formal homework given between sessions.

IPT had three phases. Phase one focussed on orientating the patient to the treatment, firstly by reviewing their depressive symptoms and interpersonal situation. Particular emphasis was placed on giving the patient the 'sick role' as a social role that engendered certain interpersonal expectations (i.e. that the patient was suffering from a genuine illness that required treating, and they needed time to recover) and absolved them from their social and occupational responsibilities. The nature and rationale of IPT was then explained. The main task of this phase was identifying and agreeing on one or two interpersonal problem areas to focus on during treatment. IPT emphasised four major interpersonal problem areas based on research into life events (Paykel et al., 1969). These were; 1) 'interpersonal role disputes' (i.e. interpersonal conflict/s, 2) 'role transitions' (i.e. changes in interpersonal roles), 3) 'grief' (i.e. interpersonal loss), and 4) 'interpersonal deficits' (i.e. limited social skills / social support) (Klerman et al., 1984). Phase two constituted the main body of treatment, and focussed on helping the patient address the interpersonal problem area/s identified in the first phase of treatment (see Table 1 for a description of the treatment

procedures associated with each problem area). Finally, phase 3 concerned ending treatment and preventing relapse.

1.2.3. Current Evidence-Base for Cognitive Therapy and Interpersonal Psychotherapy for Depression

CT and IPT for depression now have well-established evidence bases, and have been compared head to head in a number of studies (Elkin et al., 1989; Lemmens et al., 2015; Luty et al., 2007; Peeters et al., 2013; Shea et al., 1992). The best known of these is the National Institute of Mental Health Treatment of Depression Collaborative Research Program, a four arm RCT that compared CT, IPT, antidepressant medication plus clinical management (minimal supportive therapy), and placebo medication plus clinical management (Elkin et al., 1989; Elkin, Parloff, Hadley, & Autry, 1985; Shea et al., 1992). This study found similar levels of effectiveness for all treatments at the end of treatment and at 18 month follow-up (Elkin et al., 1989; Shea et al., 1992). This finding has been supported in a recent meta-analysis (N=243 studies) of RCTs comparing psychological treatments for depression (Cuijpers, Andersson, Donker, & van Straten, 2011).

CT and IPT have since been adapted for use in different cultures and in a variety of health care settings (Beck, 2005; Markowitz & Weissman, 2012). Both are currently frontline treatments for depression in the UK (National Institute for Health and Care Excellence, 2009). Both have also been modified to treat other psychological disorders (Beck, 2005; Markowitz & Weissman, 2012). Cognitive behaviour therapy for bulimia nervosa (CBT-BN) and interpersonal psychotherapy for bulimia nervosa (IPT-BN) are two such adaptations.

Table 1: Summary of interpersonal problem areas and their treatment strategies

Problem Area	Aim	Treatment Strategies
Interpersonal role disputes	To address interpersonal conflict with the patients' significant other/s	a) Increase understanding of the nature of the role dispute, particularly differing role expectations b) Explore alternative ways of managing the dispute c) Resolve the dispute (e.g. by modifying maladaptive communication styles, tackling unreasonable expectations, or ending the relationship)
Role transitions	To help the patient adjust to changes in the interpersonal role/s e.g. entering retirement, becoming a parent	a) Evaluate the significance of the former role b) Express emotions associated with the transition c) Evaluate the significance of the new role and ability to meet the demands of the new role d) Develop new social support
Grief	To facilitate the grieving process	a) Evaluate the loss and its implications b) Normalise the grieving process c) Develop new interests/relationships to replace the relationship that has been lost
Interpersonal deficits	To reduce social isolation, increase close interpersonal relationships and improve interpersonal skills	a) Evaluate past significant relationships b) Identify common problems / patterns in these relationships c) Find new ways to interact with others to reduce social isolation and improve the quality of their relationships
Life goals (IPT-ED only)	To help the patient consider their future life goals, which may have been interrupted by the eating disorder	a) Explore aspirations b) Identify potential steps to achieve aspirations c) Implement steps to achieve aspirations

Notes. IPT-ED=interpersonal psychotherapy for eating disorders (discussed later in the chapter).

1.3. Cognitive Behaviour Therapy and Interpersonal Psychotherapy for Bulimia Nervosa

CT and IPT were adapted for BN in the 1980s at the Centre for Research on Eating Disorders at Oxford (CREDO) led by Christopher Fairburn (Fairburn, 1981; Fairburn et al., 1991). At the time, BN was a new diagnostic category, although its similarity to anorexia nervosa (AN) had been documented (Russell, 1979). Whilst BN shared the morbid fear of fatness seen in AN, and similar dieting behaviour, BN was characterised by episodes of uncontrolled overeating ('binge eating') followed by compensatory behaviour such as self-induced vomiting (Russell, 1979). These forms of behaviour were an attempt to 'eliminate' from the body what had been consumed during the binge (Russell, 1979). Binge eating typically occurred in secret and involved eating unusually large amounts of food in a short space of time, and it was accompanied by a sense of loss of control over eating (Russell, 1979). The episodes were a source of distress, shame and guilt (Russell, 1979). Unlike AN, patients with BN were typically of a normal weight (Russell, 1979).

Being a new diagnosis, there had been no research into the treatment of BN and, based on his clinical experience, Russell (1979) described the disorder as "intractable" (Russell, 1979). CBT-BN (Fairburn, 1984; Fairburn, Cooper, & Cooper, 1986) was one of the first psychological treatments for this disorder. It was based on the general strategy used in CT for depression but its procedures differed markedly. The first report on the treatment was a case series (Fairburn, 1981), and it was then tested in a small RCT (Fairburn, Kirk, O'Connor, & Cooper, 1986). It was this trial that prompted the development of IPT-BN due to the unexpected success of the focal form of

supportive psychotherapy utilised as a control condition (Fairburn, 1992). This psychotherapy had much in common with IPT in terms of its structure and non-directive stance, with the core of the treatment focusing on accompanying problems and how they related to the eating disorder (Fairburn, 1992). As the majority of these were found to be interpersonal in nature, this suggested that IPT might be a suitable alternative treatment for BN (Fairburn, 1992).

1.3.1. Description of Cognitive Behaviour Therapy for Bulimia Nervosa

1.3.1.1. Theoretical Model

Similar to Beck's (1963) cognitive model of depression, the cognitive behavioural model of BN (see Figure 1) emphasised the role of cognitive processes. In particular, Fairburn and colleagues (1984; 1986) highlighted the presence of dysfunctional beliefs regarding the importance of shape and weight in how patients' evaluated their self-worth (with thin being seen as good, and fat as bad). Fairburn (1984) proposed that it was this over-evaluation of shape and weight that underlay patients' fears of fatness and ultimately drove the patient to both monitor and (attempt to) control their shape, weight and eating. Furthermore, he and his colleagues proposed that whilst it was the patients' dysfunctional beliefs regarding shape and weight that drove the disorder, it was the manner in which they implemented these behaviours that subsequently locked the patient into the disorder (Fairburn, Cooper, et al., 1986). Dieting was typically implemented in a rigid and inflexible manner, with patients imposing strict dietary rules on their eating (e.g., avoiding certain foods, not eating at certain times, setting precise limits on the amount they ate). Such rules were so inflexible as to be

practically impossible to adhere to, particularly during times of stress. 'Rule breaking' was typically reacted to in an all-or-nothing manner, with transgressions resulting in patients completely (albeit temporarily) abandoning their dietary restraint, and binge eating instead. Although binge eating might initially be experienced positively by patients as a relief from their normally rigid control over eating, and a distraction from current difficulties, these effects were short-lived (Fairburn, Cooper, et al., 1986). Binge eating was experienced as evidence of lack of control over eating, strengthening the patient's fears about becoming fat. This resulted in the patient returning to the strict dieting that had left them vulnerable to binge eating in the first place. Furthermore, such fears led the patient to utilise compensatory weight-control behaviour to 'eliminate' what had been consumed during the binge e.g. self-induced vomiting (Fairburn, Cooper, et al., 1986). Somewhat counterintuitively, the compensatory behaviours used to 'counter-act' the binge, as well as the initial 'positive' effects of binge eating, both inadvertently encouraged further episodes of binge eating, the former through a belief that binge eating was being adequately compensated for, thereby encouraging larger and more frequent bingeing, and the latter as a new way of managing difficult mood states. Fairburn and colleagues (1986) argued that treatment needed to focus on the over-evaluation of shape and weight, the way that the patient was eating, and the triggers of binge eating.

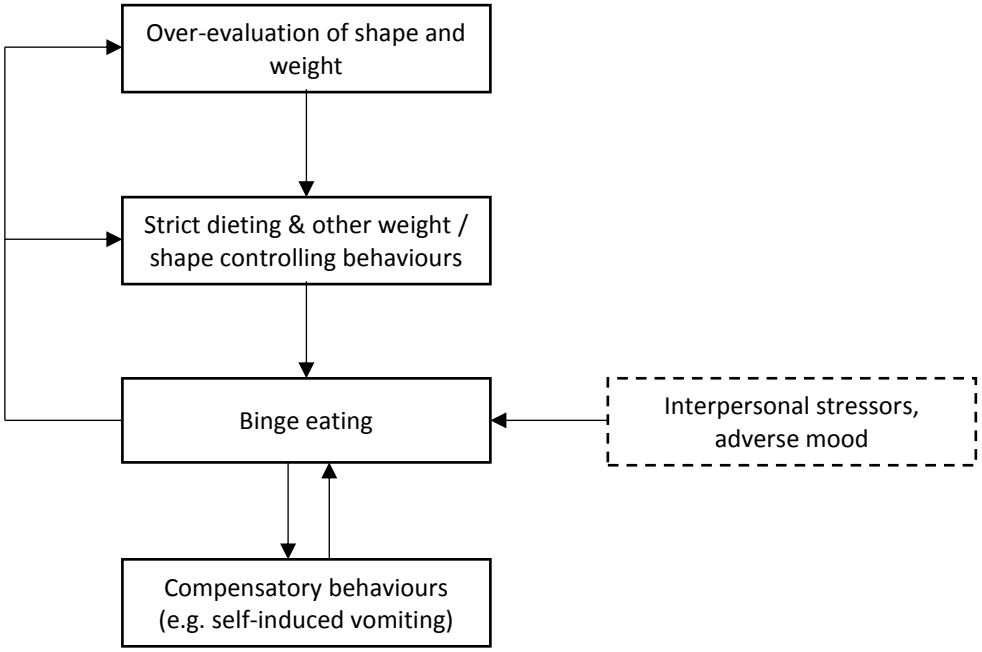


Figure 1: The cognitive behavioural model of bulimia nervosa (modified from Fairburn, Marcus, & Wilson, 1993).

1.3.1.2. The Treatment

CBT-BN consisted of a maximum of 20 sessions of 50 minutes over approximately 20 weeks (Fairburn, 1984; Fairburn, Marcus, et al., 1993). It was similar in structure to CT, with sessions occurring twice-weekly for the first four weeks of treatment, and weekly thereafter until end of treatment, when they became fortnightly, thus incorporating the "tapering off" of therapy that Beck and colleagues (Beck et al., 1984, 1979) had recommended. As in CT, 'homework' between sessions formed a key part of the treatment.

CBT-BN had three stages. Similar to CT, stage 1 focussed on explaining the rationale of treatment, providing psychoeducation regarding the disorder, creating a formulation of the patients' eating problem and achieving early behavioural change (Fairburn, 1981). There were two key behavioural interventions at this stage; the implementation of a regular eating pattern to begin tackling the tendency to diet, and

the establishment of weekly routine weighing by the patient to enable the patient to accurately monitor their weight in the face of changing eating habits (Fairburn, 1984; Fairburn, Marcus, et al., 1993). As in CT, self-monitoring was a key treatment procedure and was used throughout the duration of therapy, although in this case the monitoring focused mainly on the patient's behaviour rather than their thoughts. Interviews with significant others were also incorporated into the treatment.

Stage 2 of treatment used some cognitive restructuring as described by Beck (Beck et al., 1984, 1979) but it was largely behavioural in nature, the goal being to modify the patients' over-evaluation of shape and weight (Fairburn, 1984; Fairburn, Marcus, et al., 1993). This included addressing behaviours such as frequent shape checking. In addition, there was a continued focus on addressing dieting behaviour, particularly dietary rules related to excluded foods and limits on dietary intake. This involved supporting the patient to gradually introduce 'avoided foods' into their normal diet or slowly increasing their dietary intake so they were no longer undereating. Use of compensatory strategies (e.g. vomiting) following these changes was prevented ('exposure response prevention'). The treatment also included training in problem solving skills to help the patient deal more effectively with adverse events. Similar to CT, stage 3 of treatment focused on relapse prevention.

1.3.2. Description of Interpersonal Psychotherapy for Bulimia Nervosa

1.3.2.1. Theoretical Model

IPT-BN was based on the same theoretical rationale as IPT for depression; namely, that interpersonal problems served to trigger or maintain aspects of the disorder (Fairburn,

1992, 1997). Interpersonal functioning was often disturbed in BN, as suggested by the initial RCT of CBT-BN against supportive psychotherapy (Fairburn, Kirk, et al., 1986). Binge eating was often precipitated by negative interpersonal events (Fairburn, 1997). Whilst based on the same rationale as IPT for depression, the focus of treatment in IPT-BN concentrated solely on the current, maintaining interpersonal problems, rather than the precipitating ones (Fairburn et al., 1991). This was because BN typically had a chronic, and unremitting course, rather than the episodic pattern of depression (Fairburn et al., 1991). For many, this meant that antecedent interpersonal problems were no longer relevant in maintaining their disorder (Fairburn et al., 1991).

1.3.2.2. The Treatment

IPT-BN consisted of up to 20 sessions of treatment over a 20 week period (Fairburn, 1992, 1997). Whilst similar to IPT for depression, it was closer in structure to CBT-BN. Sessions lasted 50 minutes, rather than an hour, and occurred twice weekly for the first 4 weeks of treatment, changing to weekly, and then to fortnightly by the end of treatment (Fairburn, 1992, 1997). These changes were largely related to its development as a comparison treatment for CBT-BN with its structure adapted accordingly (Fairburn et al., 1991). Otherwise, the treatment remained similar to IPT for depression; sessions being largely unstructured, and becoming increasingly patient-led as treatment progressed (Fairburn, 1992, 1997). It had the same three phases of treatment as IPT for depression and focused on the same problem areas (see Table 1).

There were three main differences between IPT for depression and IPT-BN. Firstly, in phase one the interpersonal review focused on the patients' interpersonal context prior and subsequent to the development of BN, and current related interpersonal

difficulties (Fairburn, 1997). Secondly, the emphasis on giving the patient the 'sick role' was dropped from the treatment (C. G. Fairburn, personal communication, July 31, 2016). This was removed as it was considered that giving the patient the 'sick role' was potentially counter-productive due to the often chronic nature of eating disorders. Instead, patients were encouraged to maintain their social and occupational responsibilities, rather than reducing them. Finally, in phase two, the therapist avoided discussing the patient's symptoms in order to prevent these from dominating the session and usurping the focus on interpersonal issues (Fairburn, 1997).

1.3.3. Current Evidence-Base for Cognitive Behaviour Therapy and Interpersonal Psychotherapy for Bulimia Nervosa

CBT-BN and IPT-BN now have established evidence bases (Hay, 2013; Hay, Bacaltchuk, Stefano, & Kashyap, 2009). Unlike CT and IPT for depression, research suggests different rates of action in CBT-BN and IPT-BN (Hay et al., 2009; National Institute for Health and Care Excellence, 2004; Polnay et al., 2014; Shapiro et al., 2007). CBT-BN has a quicker effect than IPT-BN, but the difference between the two treatments disappears by 8 to 12 month follow-up (Agras, Walsh, Fairburn, Wilson, & Kraemer, 2000; Fairburn et al., 1991; Fairburn, Jones, Peveler, Hope, & O'Connor, 1993). This is due to continuing improvement following treatment in those who have received IPT-BN whereas in CBT-BN treatment outcomes are maintained from end of treatment to follow-up (Agras et al., 2000; Fairburn, Jones, et al., 1993). Many national clinical guidelines recommend CBT-BN as the first line treatment for BN with IPT-BN being an alternative option (e.g. American Psychiatric Association, 2006; Hay et al., 2014; National Institute for Health and Care Excellence, 2004).

1.4. Enhanced Cognitive Behaviour Therapy and Interpersonal Psychotherapy for Eating Disorders

CBT-BN and IPT-BN were adapted into transdiagnostic treatments in the late 1990s by the same research group that was responsible for the development of CBT and IPT for BN (CREDO). The adaptation of CBT-BN was stimulated by the recognition that AN and BN shared many characteristics (Fairburn, Cooper, & Shafran, 2003). Both had the same over-evaluation of shape and weight, and their control, and were characterised by similar behaviour, including dieting, and bingeing and purging (Fairburn et al., 2003). The main difference appeared to lie in the relative balance of overeating and undereating in the two disorders (Fairburn et al., 2003). Furthermore, it was noted that there was commonly diagnostic 'drift' between the eating disorder diagnoses, particularly from AN to BN (e.g. Agras et al., 2000; Sullivan, Bulik, Fear, & Pickering, 1998), thereby further undermining the distinction (Fairburn et al., 2003). Fairburn and colleagues (2003) hypothesised that AN, BN and the atypical eating disorders (such as 'eating disorder not otherwise specified (EDNOS)' and 'binge eating disorder' (BED)) shared common maintaining mechanisms, and therefore might respond to similar treatment approaches (Fairburn et al., 2003).

CBT-E and IPT-ED built on the treatment models of CBT-BN and IPT-BN respectively, with some adaptations to make the treatments suitable for a broader range of eating disorder presentations. However, whilst relatively few adaptations were made for IPT-ED, CBT-E itself underwent a significant redesign in an attempt to enhance the effectiveness of the treatment (Z. Cooper & Fairburn, 2011; Fairburn et al., 2003). For example, there was much greater emphasis on modifying body image and enhancing

patients' ability to deal with setbacks. Another change was the development of two different versions of the treatment; a focused version (CBT-Ef, often referred to as just 'CBT-E'), which addressed the core psychopathology of the eating disorder, and a broad version (CBT-Eb) that incorporated additional modules designed to target processes "external" to the eating disorder which were thought to impede a subset of these patients making change during treatment (Z. Cooper & Fairburn, 2011; Fairburn et al., 2003). Accordingly, in the broad version of CBT-E there were additional treatment modules on clinical perfectionism, core low self-esteem, mood intolerance and severe interpersonal difficulties (Fairburn, Cooper, Shafran, Bohn, & Hawker, 2008a).

1.4.1. Description of Enhanced Cognitive Behaviour Therapy

1.4.1.1. Theoretical Model

CBT-E was based on a transdiagnostic model of the psychopathology of eating disorders (see Figure 2) which was an extension of the theoretical model of BN, with the over-evaluation of shape and weight and their control continuing to be hypothesised to be the key maintaining process of the eating disorders (Fairburn et al., 2003). As in the cognitive behavioural model of BN, over-evaluation of shape and weight encouraged attempts to control eating, shape and weight, mostly through dieting (Fairburn et al., 2003). For those patients who were more 'successful' at maintaining their 'control', dieting led to weight loss and potentially becoming underweight. In these patients, the resultant effects of starvation then locked the patient into the eating disorder (Fairburn et al., 2003). Heightened preoccupation with food and increased feelings of fullness, reinforced the patients' concerns about losing

control over their eating and becoming fat, leading them to increase their efforts to control their eating (Fairburn et al., 2003; Keys, Brozek, Henshel, Mickelson, & Taylor, 1950). Conversely, low weight also made patients vulnerable to binge eating due to the physiological pressure to eat resulting from starvation (Fairburn et al., 2003; Keys et al., 1950). As described in the cognitive behavioural model of BN, dieting made the patient vulnerable to binge eating due to the strict way in which dieting was implemented, with the breaking of dietary rules leading to a temporary abandoning of dietary restraint (Fairburn et al., 2003). Such episodes could be triggered by adverse events and moods, hypothesised to later reinforce tendencies to binge when stressed due to the positive effects bingeing had on affect (Fairburn et al., 2003). Irrespective of the cause, binge eating episodes reinforced patients' fears of becoming fat and their need to control their eating, provoking the use of compensatory behaviour to 'counteract' the effects of the binge. Such episodes were hypothesised to be inherently self-reinforcing due to the belief that these forms of behaviour adequately compensated for what was eaten during the binge, resulting in larger or more frequent bingeing. It was hypothesised that in underweight patients, drift towards a more bulimic presentation was driven by these processes (Fairburn et al., 2003).

In addition to the core psychopathology described above, the model identified four accompanying problems that complicated the treatment of a subset of these patients. These were; 'clinical perfectionism' (an over-evaluation of achievement and achieving), 'core low self-esteem' (a pervasive, longstanding negative view of the self, unrelated to depression), 'major interpersonal problems' and 'mood intolerance' (an inability to tolerate or manage certain mood states), although the latter later became

incorporated into the focused version of treatment (Fairburn, Cooper, Shafran, et al., 2008b).

As in CBT-BN, it was hypothesised that CBT-E needed to address both the patients' over-evaluation of shape and weight and their disturbed eating habits and, if present and marked, the accompanying external processes.

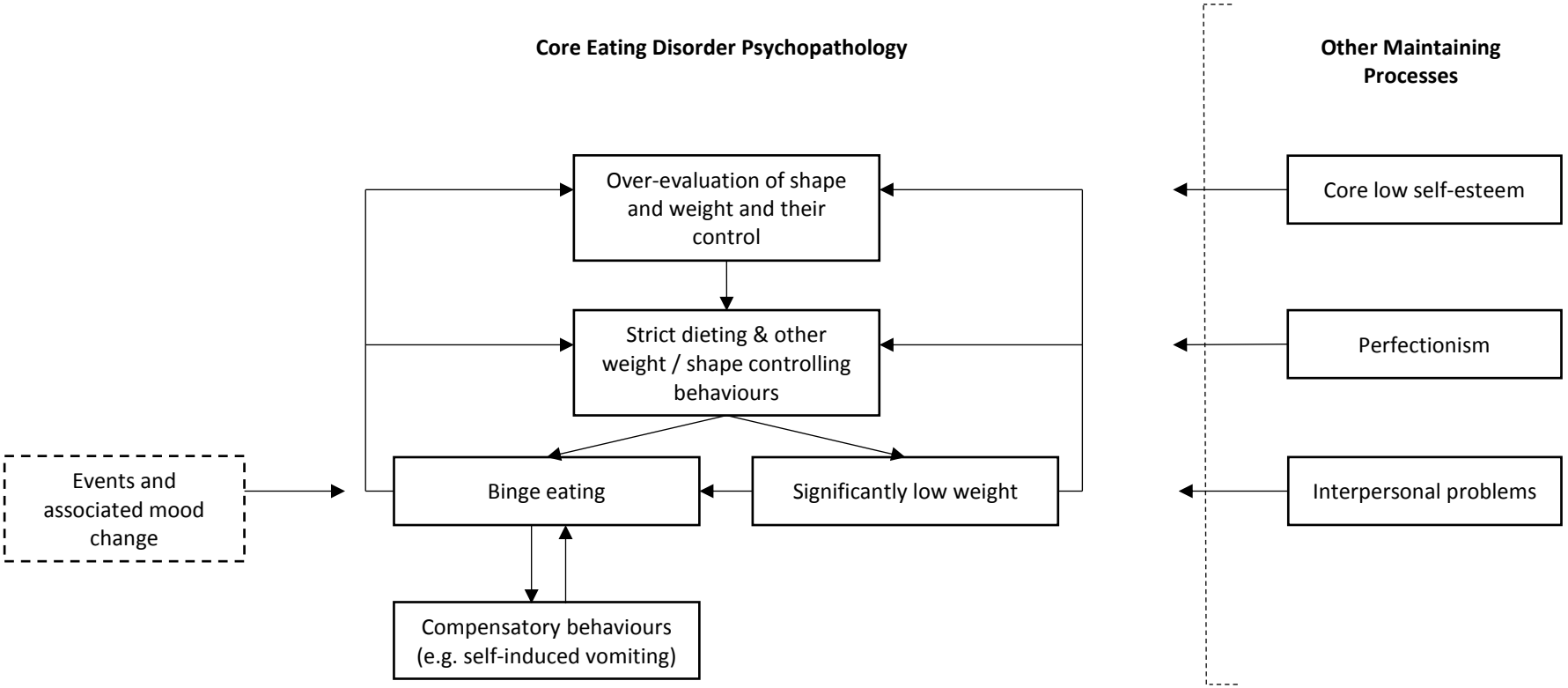


Figure 2: The enhanced cognitive behavioural model of eating disorders (modified from Fairburn (2008))

1.4.1.2. The Treatment

The structure of CBT-E (both the focused and broad versions) was similar to CBT-BN, consisting of 20 50-minute sessions delivered over 20 weeks, with session spacing changing from twice weekly, to weekly and then fortnightly by the end of treatment (Fairburn, Cooper, Shafran, et al., 2008b). There were three key changes; firstly, CBT-E had four stages of treatment, rather than three (Fairburn, Cooper, Shafran, et al., 2008b). Secondly, for patients who were underweight, treatment was extended to 40 weeks to support the patient to regain and maintain a healthier weight (Fairburn, Cooper, Shafran, et al., 2008b). Finally, the treatment incorporated a review session 20 weeks after treatment ended to check on patients' progress since finishing treatment (Fairburn, Cooper, Shafran, et al., 2008b).

As in CBT-BN, stage one of CBT-E focused on orientating the patient to the treatment and achieving early behaviour change (Fairburn, Cooper, Shafran, et al., 2008b). It included similar interventions, with 'regular eating' being identical in both treatments, and routine weighing being replaced with 'weekly weighing' in session with the therapist to facilitate more accurate interpretation of weight and give further support to patients in decreasing their frequency of weighing (Fairburn, Cooper, Shafran, et al., 2008b). Self-monitoring was implemented in this stage and continued to play a key role throughout treatment. Stage two was new. It constituted a formal review to assess the patients' progress and any barriers to treatment, to update their formulation and to decide whether the focused or broad version of the treatment was more appropriate (Fairburn, Cooper, Shafran, et al., 2008b). Stage three addressed similar issues to stage two of CBT-BN, namely, addressing over-evaluation of shape,

reducing dieting, and enhancing problem solving. These topics were covered in more depth in formalised modules that were implemented in a flexible manner (i.e. treatment was personalised) to target the core maintaining processes identified in the specific patient's formulation. CBT-E incorporated three additional treatment procedures; 'marginalised domains', which attempted to reduce the importance of shape and weight, and their control by helping the patient to increase other areas of their life which they could use to evaluate their self-worth; 'mindsets', which focused on identifying early warning signs of setbacks and helping patients address them promptly and effectively; and 'binge-analysis' which was used to help patients identify and tackle the triggers of any residual episodes of binge eating (Fairburn, Cooper, Shafran, et al., 2008b). For those patients undertaking the broad version of treatment, stage three was also when problems with clinical perfectionism, core low self-esteem and interpersonal problems were addressed (Fairburn, Cooper, Shafran, et al., 2008b). Similar to CBT-BN, the last stage of treatment, stage 4, focused on relapse prevention (Fairburn, Cooper, Shafran, et al., 2008b).

1.4.2. Description of Interpersonal Psychotherapy for Eating Disorders

1.4.2.1. Theoretical Model

IPT-ED shared the same theoretical rational as IPT-BN, emphasising the role of interpersonal problems as maintaining processes (Murphy, Straebler, Basden, Cooper, & Fairburn, 2012). However, whilst IPT-BN focused solely on current interpersonal problems as maintaining processes, IPT-ED acknowledged distinct, although interrelated, roles for both day-to-day interpersonal difficulties as well as long-

standing interpersonal problems (see Figure 3) (Murphy et al., 2012). These problems were thought to maintain the eating disorder in two ways. Firstly, they undermined self-esteem leaving the patient feeling vulnerable. It was hypothesised that the patients' focus on controlling eating, shape, and weight was a maladaptive attempt to maintain their self-esteem (Murphy et al., 2012). However, the inevitable 'failures' of control with regards their eating/shape/weight, potentially triggered by day-today interpersonal difficulties, further undermined it, creating a vicious cycle (Murphy et al., 2012). Secondly, social isolation prevented interactions with peers, impeding normalisation of eating (Murphy et al., 2012). The fact that EDs typically emerged during adolescence also had important implications for treatment. In particular, adolescence was considered a crucial period for social development when critical life choices were being made. It was hypothesised that in addition to addressing current interpersonal issues (longstanding or otherwise), treatment also needed to address these interrupted developmental tasks, if relevant (e.g. education, occupation, development of social relationships) (Murphy et al., 2012).

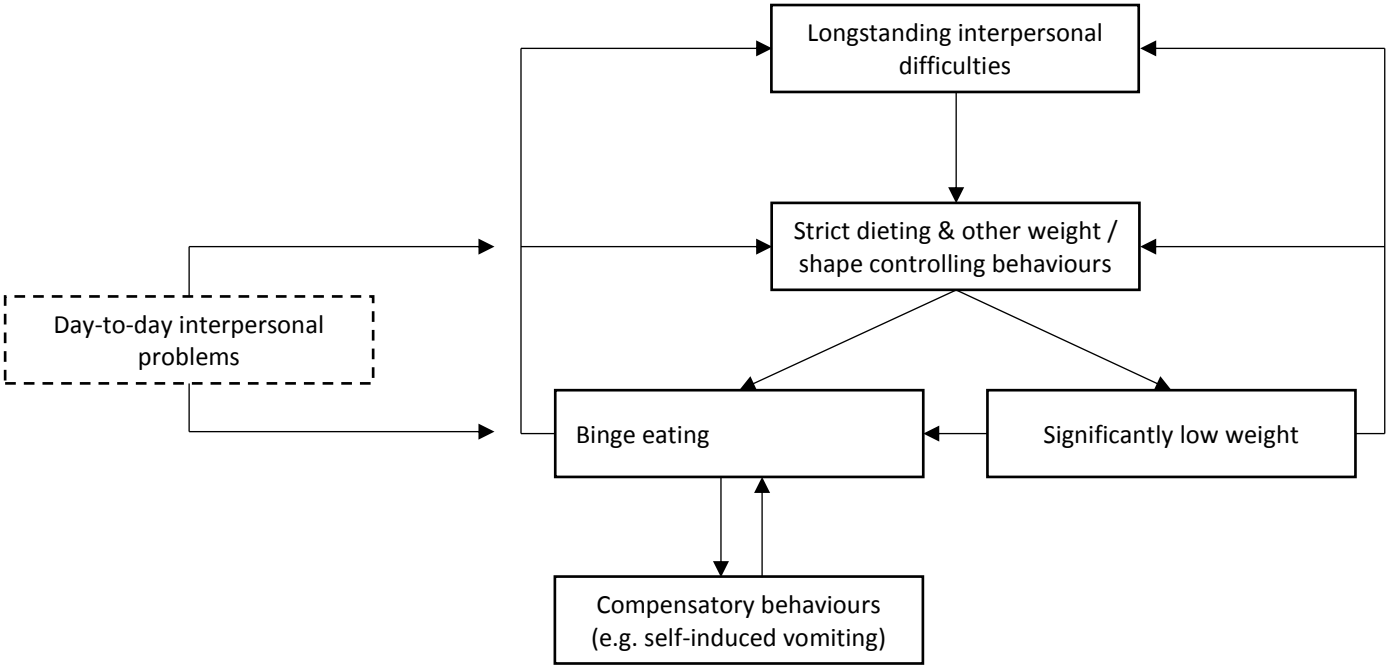


Figure 3: The interpersonal model of eating disorders

1.4.2.2. The Treatment

The treatment structure of IPT-ED was identical to IPT-BN comprising 20 50-minute sessions over a 20 week period, with sessions changing from twice weekly, to weekly and finally to fortnightly by the end of treatment (Murphy et al., 2012). Sessions remained largely unstructured, becoming increasingly patient-led as treatment progressed (Murphy et al., 2012). It also had the same three phases and goals. However, while the treatment focused on the same problem areas as IPT-BN, it included an additional problem area called 'life goals' (see Table 1) (Murphy et al., 2012). This new problem area aimed to address issues resulting from the interruption of developmental tasks associated with adolescence, encompassing a broader remit than that covered by the role transitions problem area which focused on specific changes in social roles (Murphy et al., 2012).

1.4.3. Current Evidence-Base for Enhanced Cognitive Behaviour Therapy and Interpersonal Psychotherapy for Eating Disorders

CBT-E and IPT-ED are still relatively new psychotherapies. Thus far, CBT-E has been tested in seven RCTs, as well as in eight cohort studies (see Table 2). Its focused and broad versions have been compared in three RCTS, and have been found to perform similarly (Dalle Grave, Calugi, Conti, Doll, & Fairburn, 2013; Fairburn et al., 2009; Thompson-Brenner et al., 2016), although non-underweight patients with complex psychopathology seem to do better in the broad version (Fairburn et al., 2009; Thompson-Brenner et al., 2016). The remaining RCTs have respectively compared CBT-E to integrative cognitive-affective psychotherapy (ICAT; a cognitive behavioural treatment that also addresses interpersonal and affective maintaining processes),

psychoanalytic psychotherapy, focal psychodynamic therapy, and IPT-ED. CBT-E was found to be superior to IPT-ED (Fairburn et al., 2015) and psychoanalytic psychotherapy (Poulsen et al., 2014) at both the end of treatment and at follow-up, despite the psychoanalytic psychotherapy having a much greater treatment duration (20 weeks versus two years, respectively). CBT-E and ICAT had similar treatment outcomes (Wonderlich et al., 2014), as did focal psychodynamic therapy (Zipfel et al., 2014). However, the latter RCT did not use the published manual of CBT-E and therefore may not have been true to treatment protocol. As yet there has only been one RCT comparing the effects of CBT-E for AN against other treatments (Byrne et al, personal communication), but its findings are not yet available.

Comparison of remission rates across studies is impeded by the use of different remission criteria between studies. However, two common remission criteria include; remission of psychopathology (defined as scoring below one standard deviation above community norms for eating disorder psychopathology on a clinical measure such as the Eating Disorder Examination (Fairburn, Cooper, & O'Connor, 2008)) and / or remission of behavioural symptomatology (e.g. cessation of bingeing and purging, achieving a body mass index (BMI) of 18.5 or higher). Remission rates for these criteria across these studies are summarised in Table 3 and Table 4.

The main findings are;

a) In non-underweight adult outpatients (five RCTs and three cohort studies);

a. The intent-to-treat (ITT) rates of remission of psychopathology in two large RCTs (N=130 and N=154) by CREDO were 65% at the end of treatment and 12 month follow-up (Fairburn et al., 2009, 2015). [1]
b. Rates of remission of psychopathology are lower in cohort studies embedded in routine clinical settings (40-42%). This appears to be due to higher rates of drop-out; remission of psychopathology in treatment completers is similar to rates observed in the RCTs (Byrne, Fursland, Allen, & Watson, 2011; Knott, Woodward, Hoefkens, & Limbert, 2015).
c. Rates of remission of behavioural symptomatology are lower.
d. When multiple remission criteria are applied (e.g., remission psychopathology, cessation of bingeing and purging and achieving a BMI ≥ 18.5) remission rates are even lower. For example, Turner and colleagues (Turner, Marshall, Stopa, & Waller, 2015) report rates as low as 19% across all patients (31% in treatment completers).

b) In non-underweight adolescent outpatients (one cohort study);
   a. Remission rate of psychopathology is similar to RCTs at 68% (Dalle Grave, Calugi, Sartirana, & Fairburn, 2015).

c) In underweight adult outpatients (one RCT and one cohort study);
   a. Rates of remission of psychopathology reported in the cohort study were similar to the RCTs (64%) at 60 week follow-up (Fairburn et al., 2013).

[1] It should be noted that the data from the first RCT were reanalysed for the second trial to correct for an overly conservative method of data imputation (see Fairburn et al., 2015 for more information).

b. Lower remission rates were reported in the RCT, which used a modified version of CBT-E (Zipfel et al., 2014).
c. One cohort study embedded in routine clinical settings (reported above) reports a much lower rates of remission for patients with AN. However, similar studies have found no significant differences between diagnoses (Turner et al., 2015).
d. Rates of remission of behavioural symptomatology (i.e. BMI ≥18.5) are much lower

d) In underweight adolescent outpatients (one cohort study only)
    a. Rates of remission of psychopathology in treatment completers was 97% (Dalle Grave, Calugi, Doll, & Fairburn, 2013). Overall remission rate was not reported.

e) In underweight adult and adolescent inpatients (one RCT and two cohort studies)
    a. Similar rates of remission have been found in both RCT and cohort studies
    b. Rates of remission of psychopathology were similar across RCT and cohort studies at 40-49% (Calugi, Dalle Grave, Sartirana, & Fairburn, 2015; Dalle Grave, Calugi, Conti, et al., 2013; Dalle Grave, Calugi, El Ghoch, Conti, & Fairburn, 2014).

Taken together these findings provide evidence that CBT-E can be used across all eating disorders and in adults and adolescents. When compared with other treatments it has either been superior to the comparator or of equivalent effectiveness. CBT-E is therefore a very attractive treatment for clinical services given

its clinical range and effects. Accordingly in 2015 it was strongly endorsed by the Chief Medical Officer for England who stated that it should be made available to all patients with eating disorders regardless of their age or locality (Davies, 2014, p. 11).

Table 2: Overview of research investigating the effectiveness of enhanced cognitive behaviour therapy (CBT-E) and interpersonal psychotherapy for eating disorders (IPT-ED)

Author	Year	Treatments	Treatment Context	Diagnosis	Sample	Research Design	Timing of Measures	Remission Criteria	Remission Rates in Treatment Starters (vs. Completers)
Byrne, Fursland, Allen and Watson	(2011)	CBT-E	Outpatients	EDs (BMI≥14)	Adults (N=125)	Cohort Study in Routine Clinical Setting	Baseline, EoT	Full remission: No diagnostic status, cessation of bingeing and purging and a BMI≥18.5.	EoT: 32% (vs. 56%) [AN: 18% (vs. 50%)] [BN: 45% (vs.50%)] [EDNOS: 41% (vs. 63%)]
								Full or partial remission: Meeting all but one of the above criteria	EoT: 40% (vs. 67%) [AN: 18% (vs. 50%)] [BN: 45% (vs.67%)] [EDNOS: 51% (vs. 73%)]
								Score < 1SD above community norms for eating disorder psychopathology	EoT: 42% (vs. 67%) [AN: 38% (vs. 67%)] [BN: 45% (vs.71%)] [EDNOS: 43% (vs. 63%)]
Knott, Woodward, Hoefkens and Limbert	(2015)	CBT-E	Outpatients	EDs (BMI>17.5)	Adults (N=246)	Cohort Study in Routine Clinical Setting	Baseline, EoT	Score < 1SD above community norms for eating disorder psychopathology	EoT: 40% (vs. 77%)
Turner, Marshall, Stopa and Waller	(2015)	CBT-E / CBT for eating disorders (Waller et al., 2007b)	Outpatients	EDs	Adults (N=203)	Cohort Study in Routine Clinical Setting	Baseline, EoT	Score < 1SD above community norms for eating disorder psychopathology, cessation of bingeing and purging and a BMI>18.5.	EoT: 19% (vs. 31%)

Author	Year	Treatments	Treatment Context	Diagnosis	Sample	Research Design	Timing of Measures	Remission Criteria	Remission Rates in Treatment Starters (vs. Completers)
Dalle Grave, Calugi, Sartirana and Fairburn	(2015)	CBT-E	Outpatients	EDs (BMI≥18.5)	Adolescents (N=68)	Cohort Study	Baseline, EoT	Cessation of bingeing and purging	EoT: 50% (vs. 77%)
								Score < 1SD above community norms for eating disorder psychopathology	EoT: 68% (vs. 82%)
Fairburn and colleagues	(2009)	CBT-Ef vs. CBT-Eb vs. wait-list (8 weeks)	Outpatients	EDs (BMI>17.5)	Adults (N=154)	RCT	Baseline, EoT, 20, 40 and 60 weeks FU	Score < 1SD above community norms for eating disorder psychopathology	No significant differences between treatments, however, patients with complex psychopathology did better in CBT-Eb, and vice versa. Combined rates at EoT: 51% (vs. 66%). Maintained at FU [a]
Wonderlich and colleagues	(2014)	CBT-E vs. ICAT	Outpatients	EDs (BMI≥18)	Adults (N=80)	RCT	Baseline, EoT, 4 months FU	Cessation of bingeing and purging	No significant differences between treatments. EoT: CBT-E: 23%, ICAT: 38% FU: CBT-E: 23%, ICAT: 33%
								Score < 1SD above community norms for eating disorder psychopathology	No significant differences between treatments. EoT: CBT-E: 38%, ICAT: 48% FU: CBT-E: 50%, ICAT: 55%
Fairburn and colleagues	(2015)	CBT-E vs. IPT-ED	Outpatients	EDs (BMI>17)	Adults (N=130)	RCT	Baseline, EoT, 20, 40 and 60 weeks FU	Cessation of bingeing and purging	CBT-E was superior to IPT-ED at EoT. EoT: CBT-E: 42% (vs. 42%), IPT-ED: 14% (vs. 16%) FU: CBT-E: 40% (vs.41%), IPT-ED: 39% (vs. 40%)

Author	Year	Treatments	Treatment Context	Diagnosis	Sample	Research Design	Timing of Measures	Remission Criteria	Remission Rates in Treatment Starters (vs. Completers)
								Score < 1SD above community norms for eating disorder psychopathology	CBT-E was superior to IPT-ED. EoT: CBT-E: 66% (vs. 75%), IPT-ED: 33% (vs. 38%) FU: CBT-E: 69% (vs. 70%), IPT-ED: 49% (vs.50%)
Poulsen and colleagues	(2014)	CBT-E vs. psychoanalytic psychotherapy	Outpatients	BN	Adults (N=70)	RCT	Baseline, 5, 12, 18 and 24 months	Cessation of bingeing and purging	CBT-E was superior to psychoanalytic psychotherapy. EoT: CBT-E: 42%, Psychodynamic psychotherapy: NA [b] FU: CBT-E: 66%, Psychodynamic psychotherapy: 15%
Thompson-Brenner and colleagues	(2016)	CBT-Ef vs. CBT-Eb	Outpatients	BN with BPD	Adults (N=50)	RCT	Baseline, EoT, 20 weeks and 6 months FU	Cessation of bingeing and purging	No significant differences between treatments; however, patients with complex psychopathology did better in CBT-Eb, and vice versa. Combined rates: 42% (vs. 55%). Maintained at FU [a]
Dalle Grave, Calugi, Doll, and Fairburn	(2013)	CBT-E	Outpatients	AN	Adolescents (N=49)	Cohort study	Baseline, EoT, 60 weeks FU	Weight restoration to 95% of expected weight	EoT: 28% (vs. 32%) FU: (45%)
								Score < 1SD above community norms for eating disorder psychopathology	EoT: (97%) FU: (90%)

Author	Year	Treatments	Treatment Context	Diagnosis	Sample	Research Design	Timing of Measures	Remission Criteria	Remission Rates in Treatment Starters (vs. Completers)
Fairburn, Cooper, Doll, O'Connor, Palmer and Dalle Grave	(2013)	CBT-E	Outpatients	AN	Adults (N=99)	Cohort Study	Baseline, EoT, 60 weeks FU	Score < 1SD above community norms for eating disorder psychopathology	EoT: 68% (vs. 88%) FU 64% (vs. 78%)
Zipfel and colleagues	(2014)	Modified version of CBT-E vs. focal psychodynamic therapy vs. treatment as usual	Outpatients	AN	Adults (N=242)	RCT	Baseline, 3 months after treatment started, EoT (10 months), 3 and 12 months FU	BMI>17.5 and assessor-rated diagnostic status	Focal psychodynamic therapy superior at FU. EoT: Approximately 15% in both treatments FU: Modified CBT-E: approximately 20%, focal psychodynamic therapy: approximately 35% [a]
Dalle Grave, Calugi, Ghoch, Conti, Fairburn	(2014)	CBT-E	Inpatients	AN	Adolescents (N=27)	Cohort Study	Baseline, EoT, 6 and 12 months FU	BMI≥18.5	EoT: 96% FU: 83%
								Score < 1SD above community norms for eating disorder psychopathology	EoT: 40%. Maintained at FU [a]
Calugi, Dalle Grave, Sartirana and Fairburn	(2015)	CBT-E	Inpatients	AN	Adolescents (N=46) Adults (N=49)	Cohort Study	Weekly	Weight restoration (BMI≥18.5) Weeks to restore	Significant differences between adolescents and adults. EoT: Adolescents: 65% (15 weeks), Adults: 37% (28 weeks)

Author	Year	Treatments	Treatment Context	Diagnosis	Sample	Research Design	Timing of Measures	Remission Criteria	Remission Rates in Treatment Starters (vs. Completers)
Dalle Grave, Calugi, Conti, Doll, Fairburn	(2013)	CBT-Ef vs. CBT-Eb	Inpatients	AN	Adolescents (N=23) Adults (N=57)	RCT	Baseline, EoT, 6 and 12 months FU	BMI≥18.5	No significant differences between treatments. EoT: 86% FU: 50%
								Score < 1SD above community norms for eating disorder psychopathology	No significant differences between treatments. EoT: 49%. Maintained at FU [a]

Notes. BN=bulimia nervosa, EDNOS=eating disorder not otherwise specified, AN=anorexia nervosa, RCT=randomised controlled trial, EDs=eating disorders, BN=bulimia nervosa, BMI=body mass index, EoT=end of treatment, SD=standard deviation, CBT-Ef=CBT-E focused version, CBT-Eb=CBT-E broad version, RCT=randomised controlled trial, FU=follow-up, ICAT=integrative cognitive-affective therapy, BPD=borderline personality disorder, AN=anorexia nervosa, FBT=family-based therapy. [a] Exact percentages not reported. [b] Psychoanalytic psychotherapy is a longer-term treatment than CBT-E (2 years versus 20 weeks, respectively)

Table 3: Summary of remission rates of psychopathology (score <1SD above community norm for eating disorder psychopathology) reported in studies in Table 2

Clinical Presentation	Treatment Context	Age Group	Research Design	Number of Studies	Remission Rates at End of Treatment		Remission Rates at Follow-Up	
					All Patients	**Completers Only**	**All Patients**	**Completers Only**
Non-Underweight Eating Disorders (e.g. BN / EDNOS)	Outpatients	Adults	RCT	3/5	38-65%	66-75%	50-69%	70%
			Cohort study	3/3	40-42%	67-77%	-	-
		Adolescents	RCT	0	-	-	-	-
			Cohort study	1/1	68%	82%	-	
Underweight eating disorders (e.g. AN)	Outpatients	Adults	RCT	0/1	-	-	-	-
			Cohort study	1/1	68%	88%	64%	78%
		Adolescents	RCT	0	-	-	-	-
			Cohort study	1/1	-	97%	-	90%)
	Inpatients	Adults / Adolescents	RCT	1/1	49%	-	Maintained	-
			Cohort study	1/2	40%	-	Maintained	-

Notes. BN=bulimia nervosa, EDNOS=eating disorder not otherwise specified, AN=anorexia nervosa, RCT=randomised controlled trial

Table 4: Summary of remission rates of behavioural symptomatology (cessation of bingeing and purging, body mass index ≥ 18.5) reported in studies in Table 2

Clinical Presentation	Treatment Context	Age Group	Research Design	Number of Studies	Remission Rates at End of Treatment		Remission Rates at Follow-Up	
					All Patients	**Completers Only**	**All Patients**	**Completers Only**
Non-Underweight Eating Disorders (e.g. BN / EDNOS)	Outpatients	Adults	RCT	4/5	23-42%	42-55%	23-66%	41%
			Cohort Study	2/3	19-32%	31-56%	-	-
		Adolescents	RCT	0	-	-	-	-
			Cohort study	1/1	50%	77%	-	-
Underweight eating disorders (e.g. AN)	Outpatients	Adults	RCT	1/1	15%	-	20%	-
			Cohort study	0/1	-	-	-	-
		Adolescents	RCT	0	-	-	-	-
			Cohort study	1/1	28%	32%	-	45%
	Inpatients	Adults / Adolescents	RCT	1/1	86%	-	50%	-
			Cohort study	2/2	37-96%	-	83%	-

Notes. BN=bulimia nervosa, EDNOS=eating disorder not otherwise specified, AN=anorexia nervosa, RCT=randomised controlled trial

Thus far IPT-ED has been tested in only one RCT (discussed above), which compared it to CBT-E (Fairburn et al., 2015). In this trial IPT-ED showed a similar outcome trajectory to IPT-BN, with IPT-ED showing a slower rate of response in comparison to CBT-E but with the patients continuing to improve after the end of treatment (see Figure 4). It also appeared to have a similar rate of recovery to IPT-BN, with around 50% of those who started treatment criteria for remission of psychopathology.

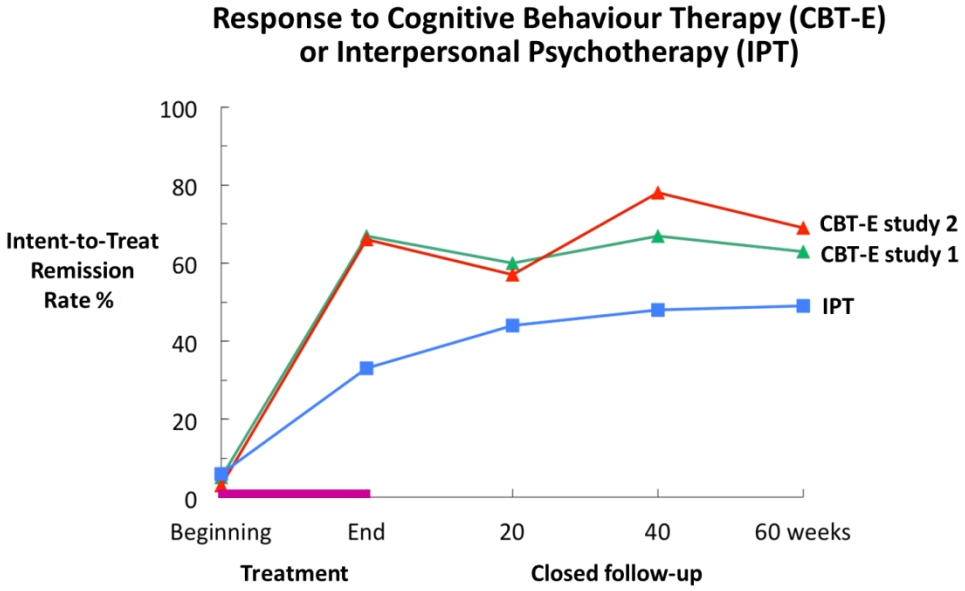


Figure 4: Comparison of remission rates in CBT-E and IPT-ED in the CREDO efficacy trials (reproduced with permission from (Fairburn et al., 2015)) [2]

There has been one two-arm RCT (N=162) comparing 20 weeks of group IPT to group CBT for obese adults (18-65 years) with binge eating disorder (BED). There were no significant differences between treatments at end of treatment or during follow-up (Hilbert et al., 2012; Wilfley et al., 2002). Rates of remission of psychopathology (defined as scoring below one standard deviation above community norms for eating disorder psychopathology in people with obesity), and cessation of binge eating were

[2] It should be noted that the data from the first RCT were reanalysed for the second trial to correct for an overly conservative method of data imputation (see Fairburn et al., 2015 for more information).

approximately 70-80% at the end of treatment, and 60% and 75%, respectively, at 12 month follow-up (Wilfley et al., 2002). At four years post-treatment the first 90 patients randomised into the RCT were followed-up. Rates of remission of psychopathology and cessation of binge eating were 55-62%, and 52-77% respectively.

There has also been one three-arm RCT (N=56) comparing 20 weeks of IPT, CBT or nonspecific supportive clinical management (i.e. minimal supportive therapy) in women with AN (17-40 years old) (Carter et al., 2011; McIntosh et al., 2005). Remission rates for psychopathology and behavioural symptoms were not reported. Overall, IPT was inferior to nonspecific supportive clinical management, but performed similarly to CBT, which performed similarly to nonspecific supportive clinical management at the end of treatment. At five year follow-up all treatments had similar outcomes (Carter et al., 2011).

1.5. Conclusion

CBT-E is the leading psychological treatment for non-underweight adults with an eating disorder, and it is one of the evidence-based treatments for those who are underweight. This accounts for approximately 75-90% and 10-25% of eating disorder cases respectively (Fairburn & Cooper, 2011; Fairburn & Harrison, 2003). IPT-ED can also be used with these patients but much less is known about its relative and absolute effectiveness. There are no competing pharmacological approaches (American Psychiatric Association, 2006; Hay et al., 2014; National Institute for Health and Care Excellence, 2004). More research is needed to establish firmly the effects of CBT-E and IPT-ED and their relative indications. Work is also needed to make them more potent.

Of relevance to such work is an understanding of how they operate. This is the focus of the following chapter.

Chapter 2. Research into the Mechanisms of Action of Enhanced Cognitive Behaviour Therapy and Interpersonal Psychotherapy for Eating Disorders

2.1. Introduction

Understanding how psychological treatments work arguably provides the strongest foundation for improving clinical treatments and patient outcomes (Kazdin & Nock, 2003). Many psychological treatments, such as cognitive behaviour therapy (CBT) and interpersonal psychotherapy (IPT), are complex interventions, particularly multi-component treatments such as enhanced CBT (CBT-E) and CBT for bulimia nervosa (CBT-BN). Without understanding a treatment's mechanisms of action (i.e. what makes it work) it is difficult to refine the treatment as it is unclear whether it is working as hypothesised, or whether some aspects are more integral than others in helping the patient to recover (Kazdin & Nock, 2003). Greater knowledge about these processes could guide treatment development and refinement through identifying which procedures are responsible for change during treatment and whether any are redundant or ineffective (Kazdin & Nock, 2003). Adaptations could then be made to make the therapy more effective by enhancing 'active' treatment procedures, and more efficient by eliminating redundant ones (Kazdin & Nock, 2003). This would have

the additional benefit of making treatments more cost-efficient (Kazdin & Weisz, 1998).

The question as to how psychological treatments work is a complex one. Treatment effectiveness is influenced by many different variables, including the characteristics of the patient, the therapist, treatment delivery, and the potency of its underlying mechanisms of action. Whilst each contributes to understanding how and when a treatment will work, the field has been hampered by a number of misconceptions regarding the role of these different variables in determining patient outcomes. One problem has been the lack of consistent use of terminology (Holmbeck, 1997; Kazdin & Nock, 2003; Kraemer, 2001).

2.1.1. Definition of Terms

2.1.1.1. Non-Specific Predictors

'Non-specific predictors' are variables that predict outcome across a variety of contexts / groups (Baron & Kenny, 1986; Dunn et al., 2015; Kraemer, Wilson, Fairburn, & Agras, 2002). In the treatment context, they predict treatment response, irrespective of the form of the treatment received (see Figure 5).

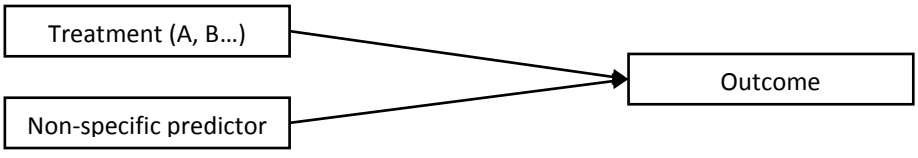


Figure 5: Illustration of the impact of a non-specific prediction on treatment outcome

2.1.1.2. Moderators

'Moderators' are variables that predict different outcomes between different groups in response to the same stimuli (Baron & Kenny, 1986; Kraemer et al., 2002). In the

treatment context, they identify subgroups that respond differently to different forms of treatments (see Figure 6). This is useful as they suggest who might respond better to one treatment over another and thus potentially help guide clinical decision making.

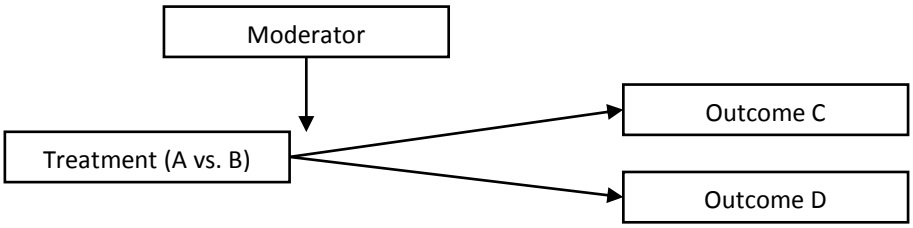


Figure 6: Illustration of the impact of a moderator on treatment outcome

2.1.1.3. Mediators

'Mediators' are variables that potentially account for the relationship between two other related variables (Baron & Kenny, 1986; Kraemer et al., 2002). In the treatment context, they are intermediary variables that account either partly ('partial mediation') or wholly ('full mediation') for the relationship between treatment and outcome (see Figure 7).

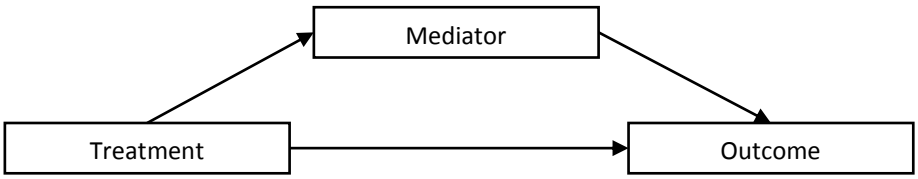


Figure 7: Illustration of the relationships between treatment, mediator and outcome

2.1.1.4. Mechanisms

It is important to make a distinction between the notions of 'mediator' and 'mechanism'. 'Mechanisms' are processes that have a *causal* role in transmitting the effect of one variable on another variable (Kraemer, Kiernan, Essex, & Kupfer, 2008; MacKinnon, Fairchild, & Fritz, 2007). In the treatment context, they are intermediary

processes that treatment changes and which influence the patient's outcome. They represent the causal mechanisms through which the treatment works. Although all mechanisms are mediators, not all mediators are causal. There is a distinction to be made between a mediator that describes a process that facilitates treatment effects ('process measures'), such as treatment adherence, and a mediator that represents the treatment's true mechanism of action (Driessen & Hollon, 2010; Kraemer & Gibbons, 2009). Kazdin and Nock (2003) proposed seven criteria that a mediator should fulfil in order to be considered a mechanism of action of treatment. These are;

1. The mediator should be strongly related to both treatment and outcome.
2. It should have a dose-response relationship to the outcome.
3. Change in the mediator should occur before change in the outcome.
4. Experimental manipulation of the mediator should result in change in the outcome
5. The mediator and outcome should have a unique, specific relationship. Multiple mediators provide a much weaker argument for one particular mediator being the mechanism of action of the treatment.
6. Research should consistently support a relationship between the mediator and the outcome.
7. There should be an empirically supported theoretical rationale for the mediator being a mechanism and this should be well-integrated within the existing evidence-base.

Whilst these criteria can be used to identify potential mechanisms of action of treatments, they are not definitive (Kazdin, 2007; Kazdin & Nock, 2003). In complex treatments it is unlikely that only one mediator exists, thus making it difficult to

disentangle treatment processes. Furthermore, the presence of non-linear relationships and moderators can present an inconsistent picture of the relationships between treatment, the mediator and the outcome. To address some of these issues Kazdin and Nock (2003) propose that the decision regarding whether a variable is a mechanism of action should be based on an accumulation of research that consistently suggests a potential causal relationship between the purported mechanism and treatment outcome.

2.1.2. Research Designs Used to Identify Mechanisms of Treatment Action

Another matter to be considered is the relative weight that should be given to the different research designs used to identify treatment mechanisms. Four research designs are discussed below with regard their strengths and weaknesses.

2.1.2.1. Observational Studies

Observational research (e.g. cohort studies) can play a role in identifying possible mechanisms through which treatments operate, for example, by identifying the timing and order of treatment effects. However, such studies are merely descriptive. It is not possible to determine whether any putative processes identified have a causal relationship with treatment outcome. They therefore provide weak evidence supporting potential mechanisms of treatment.

2.1.2.2. Experimental Studies

2.1.2.2.1. Experimental Manipulation of the Purported Mediator

Experimental manipulation of the purported mechanism is the surest way to assess its effect on the outcome. Whilst this directly tests the causal relationship between the purported mediator and the outcome, such tests are difficult to devise in the context of psychological treatment research given the number of processes operating simultaneously, or in sequence, and the clinical constraints of treatment research.

2.1.2.2.2. Dismantling Studies

Dismantling studies compare full or partial versions of the same treatment with versions that are missing one or more treatment components to assess their contribution to outcome. Such studies have the benefit of taking place within the treatment setting and being able to isolate the effects of the particular treatment components. However, they do not examine the influence of the purported mechanism within the whole treatment.

2.1.2.2.3. Efficacy and Mechanism Evaluation Randomised Controlled Trials

It has been argued that RCTs are underused for assessing how treatments work (Dunn et al., 2015; Kraemer et al., 2002). If they are to be used for this purpose they need to be designed with this in mind by embedding mediation studies within the RCT that measure the level and timing of change in both the purported mediator and the outcome (Dunn et al., 2015; Kraemer et al., 2002). They also include studies that investigate causes of different responses to treatment (National Institute for Health Research, 2017). These are known as 'efficacy and mechanism evaluation (EME)' RCTs.

Such studies are less controlled than experimental and dismantling studies, unless they employ double randomisation of both treatment and the level of the purported mediator. However, double randomisation is difficult to apply when testing a treatment (MacKinnon et al., 2007). Nonetheless, they have the potential to improve understanding of how treatments work by examining the processes of change during treatment. Furthermore this kind of research has greater ecological validity than both dismantling studies and experimental research as it directly observes the processes of interest (i.e. changes in the purported mediator and the outcome) in the context of interest (i.e. treatment).

2.2. Non-Specific Predictors of Treatment Outcome in CBT and IPT in Eating Disorders

Identifying reliable non-specific predictors of treatment outcome in eating disorder (ED) patients has been challenging. Whilst several potential non-specific predictors of treatment outcome have been identified, they have not always been reliably identified across different studies (Iacovino, Gredysa, Altman, & Wilfley, 2012; Linardon, de la Piedad Garcia, & Brennan, 2017; National Collaborating Centre for Mental Health, 2004; Steinhausen & Weber, 2009; Thompson-Brenner et al., 2013; Vall & Wade, 2015). Research has been limited by the use of small sample sizes, short lengths of follow-up, and a lack of consensus regarding the definition of outcome and what predictor variables to test (Steinhausen & Weber, 2009).

The most recent meta-analysis in this area [N=126 studies] attempted to overcome these shortcomings by combining related variables into one construct to permit

analysis of all available data (Vall & Wade, 2015). This meta-analysis identified 13 weak/moderate non-specific predictors of treatment outcome in eating disorders. Eight of these predicted poorer outcome at both end of treatment and follow-up (see Table 5).

Table 5: Non-specific predictors of poorer treatment outcome in eating disorders identified in the Vall and Wade (2015) meta-analysis

Non-Specific Predictors	Outcome at End of Treatment	Outcome at Follow-Up
Frequent bingeing and purging	X	X
Low motivation	X	X
High shape/weight concern	X	X
High levels of depression	X	X
Poor interpersonal functioning	X	X
Long eating disorder duration/early age of onset	X	X
Low body mass index (BMI)	X	X
High levels of family problems	X	X
Higher levels of eating disorder psychopathology	X	
Psychiatric comorbidity	X	
More frequent exercising	X	
Comorbid personality disorder		X
Low self-esteem		X

There have been two further studies since the publication of this meta-analysis. The first, an RCT (N=80) comparing integrative cognitive-affective therapy (ICAT) and CBT-E for bulimia nervosa (BN), found that greater dietary restraint, weight and shape concern predicted greater ED improvement at 4-month follow-up, although it also found that high levels of depression predicted worse outcome (Accurso et al., 2016). The second study, an RCT (N=130) comparing CBT-E and IPT for eating disorders (IPT-ED) in a non-underweight transdiagnostic sample, found that the over-evaluation of shape and greater duration of eating disorder were significant non-specific predictors of treatment response at 12 month follow-up (Z. Cooper et al., 2016). No other statistically significant predictors were identified. These findings are inconsistent with

those from the Accurso (2016) study (which had a markedly shorter follow-up period), but they are consistent with the meta-analysis (Vall & Wade, 2015).

2.3. Moderators of Treatment Outcome in CBT and IPT in Eating Disorders

Only two RCTs have examined potential moderators of response to CBT and IPT in eating disorders (Z. Cooper et al., 2016; Hilbert et al., 2007). The first RCT (N=162), which compared group CBT to group IPT for binge eating disorder (BED), found no treatment moderators, even when assessing variables that might theoretically be thought to differentiate the two treatments such as social functioning (Hilbert et al., 2007). The second RCT (N=130), which compared CBT-E to IPT-ED in a non-underweight transdiagnostic sample found only one candidate moderator; patients with low self-esteem appeared to do better in CBT-E compared to IPT-ED (Z. Cooper et al., 2016). Although significant, the confidence interval for this effect was wide. The result should also be interpreted in the context of IPT-ED generally performing less well than CBT-E (see Fairburn et al., 2015).

2.4. Mechanisms of Action of CBT and IPT in Eating Disorders

Research comparing CBT and IPT in eating disorders suggests that the two treatments have different outcome trajectories, at least in patients with bulimia nervosa and atypical eating disorders (Fairburn et al., 1991, 2015; Fairburn, Jones, et al., 1993). This suggests that IPT operates differently to CBT, possibly via a route that takes longer to affect the outcome (e.g. via interpersonal change). Although this difference is not

seen in binge eating disorder (Wilfley et al., 2002), this may be due to its greater treatment responsivity in general (National Collaborating Centre for Mental Health, 2004). There is also evidence to suggest that early treatment response is important in determining treatment outcome (Linardon, Brennan, & de la Piedad Garcia, 2016; Linardon et al., 2017; Nazar et al., 2017). The majority of this research has been in CBT, although what little research has been done with other treatments (including IPT) suggests it may also be important in other treatments (Linardon et al., 2016). Research into treatment-specific purported mechanisms of CBT and IPT are examined below.

2.4.1. Observational Studies

There have been four cohort studies (Calugi, El Ghoch, & Dalle Grave, 2017; Legenbauer, Schütt-Strömel, Hiller, & Vocks, 2011; Spangler, Baldwin, & Agras, 2004; Zendegui, West, & Zandberg, 2014) and two qualitative studies (Onslow, Woodward, Hoefkens, & Waddington, 2016; Waller, Evans, & Pugh, 2013) investigating purported mechanisms of CBT in eating disorders (see Table 6 for details). One of these (Spangler et al., 2004) was embedded within a larger EME RCT, which will be discussed later.

Two out of two cohort studies have supported a positive relationship between dietary restraint / eating pattern and frequency of bingeing and vomiting (Spangler et al., 2004; Zendegui et al., 2014). This was also supported in one of the qualitative studies (Waller et al., 2013). Dietary restraint was also found to be positively related to weight concern (Spangler et al., 2004). Evidence has been more mixed for the role of over-evaluation of shape, weight and control over eating. One cohort study found that dysfunctional beliefs were positively related to dietary restraint and frequency of binge

eating (Legenbauer et al., 2011). However, this was not supported in another study (Spangler et al., 2004). Likewise, a positive relationship between dysfunctional beliefs and weight concern, but not shape concern, was found in one study (Spangler et al., 2004). Another cohort study found a positive relationship between change in body checking behaviour and change in eating disorder psychopathology and weight concern (Calugi, Ghoch, & Grave, 2017).

The Spangler and colleagues study, also investigated the relationships between treatment personalisation using therapists' ratings of their use of cognitive and behavioural techniques in the treatment session. There was limited support for a direct relationship between use of these strategies and purported mediators and the outcome. The exception being a relationship between the use of behavioural techniques, body-related dysfunctional beliefs and weight concern.

There have been no observational studies of IPT.

2.4.2. Experimental Studies

2.4.2.1. Experimental Manipulation of the Purported Mediator

There have been no within-treatment studies that have attempted to manipulate purported mediators of CBT or IPT in the context of treatment.

2.4.2.2. Dismantling Studies

There have been seven dismantling studies (Agras, Schneider, Arnow, Raeburn, & Telch, 1989a; P. J. Cooper & Steere, 1995; Fairburn et al., 1991; Fairburn, Jones, et al., 1993; Freeman, Barry, Dunkeld-Turnbull, & Henderson, 1988; Kirkley, Schneider, Agras, & Bachman, 1985; Thackwray, Smith, Bodfish, & Meyers, 1993; Wilson, Rossiter,

Kleifield, & Lindholm, 1986) comparing full or partial versions of CBT with versions that are missing one or more treatment components (see Table 6 for details). These studies have found that CBT is superior to its basic components such as self-monitoring and psychoeducation (Freeman et al., 1988; Kirkley et al., 1985; Leitenberg & Rosen, 1989; Thackwray et al., 1993). The majority of studies have found that CBT is superior to versions including only its behavioural components (Fairburn et al., 1991; Fairburn, Jones, et al., 1993; Thackwray et al., 1993). Cognitive restructuring is also superior to behavioural components alone (P. J. Cooper & Steere, 1995). This is mainly due to a greater rate of relapse when cognitive elements are excluded. The one study that found these versions to be more or less equivalent had not completed follow-up assessments (Freeman et al., 1988). Similarly, versions of CBT including in-session exposure response prevention have been found superior to cognitive restructuring alone (Wilson et al., 1986), although a comparison of CBT to CBT with in-session exposure response prevention still found CBT to be superior (Agras et al., 1989a).

In addition, there have been three dismantling studies (Dalle Grave, Calugi, Conti, et al., 2013; Fairburn et al., 2009; Thompson-Brenner et al., 2016) comparing CBT-E (focused version; CBT-Ef) to CBT-E (broad version; CBT-Eb) (see Table 6 for details). In all these studies CBT-Ef and CBT-Eb have performed equivalently. However, in two of these studies, complex psychopathology has emerged as a potential moderator, with patients with complex psychopathology doing better in CBT-Eb (Fairburn et al., 2009; Thompson-Brenner et al., 2016). This finding was not replicated in an inpatient anorexia nervosa sample (Dalle Grave, Calugi, Conti, et al., 2013). However, it was speculated that the external maintaining processes addressed in CBT-Eb were less

relevant at a low-weight or in the artificial environment of the inpatient ward (Dalle Grave, Calugi, Conti, et al., 2013).

There have been no dismantling studies of IPT. It has been argued that IPT is too cohesive a treatment to be dismantled (Murphy et al., 2009). Although this view is not without dispute (Lipsitz & Markowitz, 2013).

2.4.2.3. Efficacy and Mechanism Evaluation Trials

There have been three EME studies (Ellison et al., 2016; Shah, Passi, Bryson, & Agras, 2005; Wilson, Fairburn, Agras, Walsh, & Kraemer, 2002) (see Table 6 for details). Two of these studies come from the same RCT by Agras and colleagues (2000). All three studies supported a relationship between addressing dietary restraint (in particular, changing eating pattern) and reducing frequency of bingeing and purging in CBT. Although additional purported mediators of both CBT and IPT were examined in the Wilson (2002) RCT, which compared purported mediators of CBT-BN and IPT-BN, only dietary restraint and increased self-efficacy for resisting food and eating cues to binge eat explained differing response between these treatments. Change in shape and weight concerns, interpersonal problems, self-esteem and other forms of self-efficacy was similar between treatments, and their relationship with treatment outcome was not explored.

There have been no other EME trials of IPT.

Table 6: Overview of treatment research investigating treatment-specific purported mechanisms of cognitive behaviour therapy (CBT) and interpersonal psychotherapy (IPT) in eating disorders

Author	Year	Treatment	Sample	Research Design	Timing of Measures	Treatment Variable	Purported Mediator/s	Outcome/s	Results
Spangler, Baldwin, & Agras	(2004)	CBT-BN (Fairburn, Marcus, et al., 1993)	Patients with BN (N=56)	Cohort study embedded in an RCT [a]	Bi-weekly	Likert ratings of the use of behavioural and cognitive techniques in each treatment session	Dietary restraint Body-related dysfunctional beliefs	Frequency of vomiting Weight concern Shape concern	Only behavioural techniques were positively associated with decreases in a purported mediator (body related dysfunctional beliefs) and an outcome (weight concern). Only decreased dietary restraint (purported mediator) was associated with decreased frequency of vomiting (outcome). Both purported mediators were associated with decreased weight concern (outcome). Neither purported mediator was related to shape concern (outcome)
Legenbauer, Schütt-Strömel, Hiller, & Vocks	(2011)	10 week group CBT-based body image intervention	Patients with EDs (N=41)	Cohort study	Baseline, EoT	NA	Dysfunctional beliefs about the body Dysfunctional beliefs about eating	Frequency of binge eating Dietary restraint	Change in dysfunctional thoughts about the body (purported mediator) was associated with change in frequency of binge eating (outcome). Change in dysfunctional beliefs about dieting (purported mediator) was associated with change in dietary restraint (outcome).

Author	Year	Treatment	Sample	Research Design	Timing of Measures	Treatment Variable	Purported Mediator/s	Outcome/s	Results
Zendegui, West, & Zandberg	(2014)	10 weeks of guided self-help CBT (Fairburn, 1995)	Students with binge eating (N=38)	Cohort study	Weekly	NA	Adherence to a regular eating pattern	Frequency of binge eating	Greater adherence to a regular eating pattern (purported mediator) was associated with decreased frequency of binge eating (outcome)
Calugi, Ghoch, & Grave	(2017)	CBT-E for inpatients	Patients with AN	Cohort study	Baseline, EoT, 6 months and 12 months FU	NA	Body checking behaviour	Eating disorder psychopathology Weight concern Shape concern	Change in body checking behaviour (purported mediator) was associated with change in eating disorder psychopathology (outcome) and weight concern.
Waller, Evans, & Pugh	(2013)	CBT for eating disorders (Waller et al., 2007a)	Patients with EDs (N=28)	Qualitative study	One-off interview in mid-late treatment (sessions 10-27)	CBT	Changing eating pattern	NA	Patients reported that changing their eating pattern (purported mediator); a) normalised their eating (e.g. prevented binge eating), b) decreased their concerns about eating (e.g. weight gain), c) increased their self-esteem and d) improved their body image.

Author	Year	Treatment	Sample	Research Design	Timing of Measures	Treatment Variable	Purported Mediator/s	Outcome/s	Results
Onslow, Woodward, Hoefkens, & Waddington	(2016)	CBT-E (Fairburn, 2008)	Patients with BN (N=8)	Qualitative study	One-off interview 1-2 years post-treatment	CBT-E	NA	NA	Helpful aspects of CBT-E reported by patients were; a) increasing understanding of eating disorder triggers and maintaining cycles (through self-monitoring and formulation), b) learning to challenge negative thoughts and improve coping (e.g. dealing with 'feeling fat', distraction). Some patients felt more in depth cognitive work was needed.
Kirkley, Schneider, Agras, & Bachman	(1985)	16 week group CBT (90 minute sessions)	Patients with BN (N=28)	Dismantling study (RCT)	Baseline, EoT, 3 months FU	Self-monitoring and psycho-education vs. CBT	NA	Frequency of bingeing and vomiting	CBT was superior to self-monitoring and psychoeducation alone at EoT, but similar at 3 month FU.
Wilson, Rossiter, Kleifield, & Lindholm	(1986)	16 week group CBT (90 minute sessions)	Patients with BN (N=17)	Dismantling study (RCT)	Baseline, EoT	Cognitive restructuring vs. in-session ERP [b] plus cognitive restructuring	NA	Frequency and cessation of bingeing and vomiting	ERP plus cognitive restructuring had superior cessation rates to cognitive restructuring alone

Author	Year	Treatment	Sample	Research Design	Timing of Measures	Treatment Variable	Purported Mediator/s	Outcome/s	Results
Freeman, Barry, Henderson, & Freeman	(1988)	15 week CBT	Patients with BN (N=112)	Dismantling study (RCT)	Baseline, 8 weeks, EoT	Waitlist vs. group psycho-education vs. BT vs. CBT	NA	Bulimic behaviours and symptoms Body dissatisfaction	All treatments were superior to wait-list. BT and CBT were superior to group psychoeducation on both outcomes at mid-treatment, but only BT remained was superior at EoT. BT and CBT performed equivalently at mid-treatment and EoT, although BT had lower body dissatisfaction at EoT.
Agras, Schneider, Arnow, Raeburn, & Telch	(1989b)	14 week CBT	Patients with BN (N=77)	Dismantling study (RCT)	Baseline, 6 weeks, EoT, 6 months FU	Wait-list vs. self-monitoring vs. CBT vs. CBT plus in-session ERP [b]	NA	Frequency and cessation of vomiting	CBT had superior cessation rates to all other treatments, including CBT plus in-session ERP at both EoT and FU.

Author	Year	Treatment	Sample	Research Design	Timing of Measures	Treatment Variable	Purported Mediator/s	Outcome/s	Results
Fairburn, Jones, Peveler, Hope, & O'Connor	(1991; 1993)	CBT-BN (Fairburn, Cooper, et al., 1986) IPT-BN (Fairburn, 1992)	Patients with BN (N=75)	Dismantling study (RCT)	Baseline, EoT, 4, 8 and 12 months FU	BT vs. CBT-BN vs. IPT-BN	NA	Eating disorder psychopathology Frequency and cessation of bingeing & purging Shape concern Weight concern Social Adjustment	BT, CBT-BN and IPT-BN had similar binge eating rates and improvements in social adjustment at EoT. CBT-BN and BT were superior to IPT-BN in reducing frequency of vomiting. CBT-BN was superior to both treatments in reducing dietary restraint and weight concern, and superior to BT in reducing shape concern. At FU, CBT had superior rates of cessation than BT, which was explained by a higher rate of relapse in BT. CBT-BN and IPT-BN had similar rates of cessation, reductions in dietary restraint, shape/weight concerns, and improvements in social adjustment.
Thackwray, Smith, Bodfish, & Meyers	(1993)	8 week condensed version of CBT-BN	Patients with BN (N=47)	Dismantling study (RCT)	Baseline, EoT, 6 months FU	Self-monitoring vs. BT vs. CBT	NA	Frequency cessation of bingeing and purging Drive for thinness	At EoT CBT and BT had superior cessation rates to self-monitoring, and BT had superior reduction of drive for thinness compared to both treatments. At FU, CBT was superior to both BT and self-monitoring, due to greater relapse in BT.

Author	Year	Treatment	Sample	Research Design	Timing of Measures	Treatment Variable	Purported Mediator/s	Outcome/s	Results
Cooper & Steere	(1995)	CBT-BN	Patients with BN (N=31)	Dismantling study (RCT)	Baseline, 4 weeks, EoT, 12 months FU	BT vs. cognitive restructuring	NA	Frequency of bingeing and vomiting	BT and CT performed equivalently at EoT. However, CBT-BN (minus ERP) was superior at FU due to greater relapse in the ERP.
Fairburn and colleagues	(2009)	CBT-E	Patients with EDs (N=154)	Dismantling study (RCT)	Baseline, EoT, 20, 40 and 60 weeks FU	CBT-Ef vs. CBT-Eb	NA	Eating disorder psychopathology	Both treatments performed equivalently at EoT and FU. However, results suggested that patients with complex psychopathology did better in CBT-Eb (potential moderator)
Dalle Grave, Calugi, Conti, Doll, & Fairburn	(2013)	CBT-E	Patients with AN (N=80)	Dismantling study (RCT)	Baseline, EoT, 6 and 12 months FU	CBT-Ef vs. CBT-Eb for inpatients	NA	Body mass index Eating disorder psychopathology	Both treatments performed equivalently at EoT. There was no indication that complex psychopathology was a potential moderator.
Thompson-Brenner and colleagues	(2016)	CBT-E	Patients with BN and BPD (N=50)	Dismantling study (RCT)	Baseline, EoT, 20 weeks and 6 months FU	CBT-Ef vs. CBT-Eb	NA	Frequency and cessation of bingeing and purging Eating disorder psychopathology	Both treatments performed equivalently at EoT. Patients with complex psychopathology did better in CBT-Eb (potential moderator)

Author	Year	Treatment	Sample	Research Design	Timing of Measures	Treatment Variable	Purported Mediator/s	Outcome/s	Results
Wilson, Fairburn, Agras, Walsh, & Kraemer	(2002)	CBT-BN vs. IPT-BN	Patients with BN (N=220)	EME trial embedded in an RCT [a]	Baseline, 2, 4, 6 and 10 weeks, EoT	CBT-BN vs. IPT-BN	Dietary restraint Weight and shape concerns Self-efficacy to resist binge eating triggers Interpersonal problems Self-esteem	Frequency of bingeing and vomiting	CBT-BN had superior outcome/s to IPT-BN at EoT but treatments performed equivalently at FU. CBT-BN was superior to IPT-BN in decreasing dietary restraint and self-efficacy for resisting binge eating triggers (purported mediators). Change in dietary restraint and self-efficacy to resist food/eating triggers at weeks 4 and 6 were associated with reduced frequency of bingeing and purging (outcome) at EoT (partial mediators). Only dietary restraint was associated with decreased frequency of bingeing at FU (full mediator).
Shah, Passi, Bryson, & Agras	(2005)	CBT-BN vs. IPT-BN	Patients with BN (N=158)	EME trial embedded in an RCT [a]	EoT	CBT-BN vs. IPT-BN	Eating pattern	Frequency of bingeing and purging	Patients who had a more regular eating pattern (purported mediator) were more likely to not binge and purge (outcome).

Author	Year	Treatment	Sample	Research Design	Timing of Measures	Treatment Variable	Purported Mediator/s	Outcome/s	Results
Ellison and colleagues	(2016)	CBT-E vs. ICAT	Patients with typical or atypical BN (N=80)	EME trial embedded in an RCT	Baseline, 4 months FU	CBT-E vs. ICAT	Eating pattern	Frequency of bingeing and purging	Increased consumption of the evening meal (purported mediator) was associated with a decrease in the outcome/s. Increased consumption of lunch (purported) was associated with increased frequency of purging (outcome).

Notes. CBT-BN=CBT for bulimia nervosa, BN=bulimia nervosa, RCT=randomised controlled trial, CBT=cognitive behaviour therapy, EDs=eating disorders, EoT=end of treatment, NA=not applicable, CBT-E=enhanced cognitive behaviour therapy, AN=anorexia nervosa, FU=follow-up, ERP=exposure response prevention, BT=behaviour therapy, IPT-BN=IPT for bulimia nervosa, CBT-Ef=CBT-E focused version, CBT-Eb=CBT-E broad version, BPD=borderline personality disorder, EME=efficacy and mechanism evaluation, ICAT= integrative cognitive-affective therapy. [a] These studies were embedded within the same RCT (Agras et al., 2000)

2.5. Conclusion

Relatively few studies have examined the purported mediators of CBT and IPT. Of those identified, the majority have focused on purported mediators of CBT. Although dismantling studies are supportive of the importance of addressing over-evaluation of shape and weight in maintaining treatment outcomes in the long-term, there has been little research examining how it changes during treatment. Furthermore, there has been no research into the impact of specific treatment procedures such as addressing frequent weighing and shape checking behaviours, despite these being a core procedures in CBT-BN and CBT-E. Whilst the relationship between dieting and binge eating has been the most researched, and has been consistently supported across studies, studies have been limited in a number of ways. The majority of research has been undertaken with patients with bulimia nervosa, and have generally used small samples sizes. There has also been an over-reliance on cohort designs. Only one study has examined purported mediators of IPT. However, as change in interpersonal problems was similar in both treatments, the relationship between interpersonal problems as a purported mediator of CBT and/or IPT was not investigated further. Other common limitations of research-to-date have been a lack of consideration of the timing and course of change in the purported mediator/s and the outcome/s, and imprecise tracking of the implementation of specific treatment procedures where treatment personalisation is an issue. There is therefore a need for more detailed, longitudinal work examining the relationships between different treatment components, their purported mediators and treatment outcomes in the context of treatment. EME trials are ideally suited for such research, but have so far been

underused. This is the design used in the studies described in subsequent chapters in this thesis.

Chapter 3. Research Design

3.1. Introduction

Randomised controlled trials (RCTs) represent an untapped resource for studying mediational processes in the context of treatment (Dunn et al., 2015; Kraemer et al., 2002). The current chapter describes an RCT comparing enhanced cognitive behaviour therapy (CBT-E) and interpersonal psychotherapy for eating disorders (IPT-ED). This trial was designed so that a series of longitudinal mediation studies were embedded in the main treatment efficacy trial (Fairburn et al., 2015; Murphy et al., 2009). Three papers have been published from this trial. These are a pre-paper describing the planned design of the mediation studies (Murphy et al., 2009), a paper on the results of the main trial outcomes (in which the current author had a minor role) (Fairburn et al., 2015), and a paper on predictors and moderators (Z. Cooper et al., 2016). Full text articles of the open access papers reporting the results of the main trial and the pre-paper for the mediation studies can be found in Appendix 1 and Appendix 2.

3.2. Overview of Main Efficacy Trial

3.2.1. Research Design and Procedure

3.2.1.1. Design

The study was a stratified RCT comparing CBT-E (Fairburn, Cooper, Shafran, et al., 2008b) and IPT-ED (Murphy et al., 2012) in a non-underweight transdiagnostic sample

of eating disorders. Minimisation was used to allocate patients to the trial based on eating disorder diagnosis, gender, and the need to continue taking psychotropic medication (e.g. antidepressant medication) during the treatment trial. Ethical approval for the study was obtained from the National Research Ethics Service Oxfordshire Research Ethics Committee C (REF: 06/Q1606/82).

3.2.1.2. Recruitment

One hundred and thirty patients were randomised from consecutive referrals to a local eating disorder clinic in central Oxfordshire between 2006 and 2011.

3.2.1.2.1. Inclusion / Exclusion Criteria

Patients had to be between 18 and 65 years, and able to give informed consent. Patients were excluded if they had a body mass index (BMI) of 17.5 or lower, they had a BMI greater than 39.9, they had previously received a treatment similar to CBT-E or IPT-ED, they had a comorbid general psychiatric disorder that prevented eating disorder-focused treatment, they were medically unstable, they were pregnant or they were not available to attend treatment. Patients undertaking psychiatric treatment not resembling CBT-E or IPT-ED gradually stopped their treatment prior to commencing the current study except for patients taking antidepressant medication that was clinically warranted. These patients were kept on a stable dose of medication throughout the trial.

3.2.1.2.2. Description of Sample

The final sample was predominantly female (N=127, 98%), white (N=124, 95%), single (N=110, 85%) and had a mean age of 25.9 years (SD=8.7). Half the sample were

students (N=65, 50%). Patients had a mean eating disorder duration of 9.9 years (SD=8.6) and just under a third of the sample had a history of anorexia nervosa (N=36, 28%). In total, 53 patients (41%) met the criteria for bulimia nervosa (BN) and 77 patients (59%) met the criteria for eating disorder not otherwise specified (EDNOS), of whom 8 patients (10%) met the criteria for binge eating disorder (BED). Just over one third had an axis I comorbidity (N=47, 36%). Fifty-seven patients (44%) remained on psychotropic medication during the trial.

3.2.1.3. Procedure

Patients were screened following referral, and if they met eligibility criteria, were consented to take part in the research. They were then randomised to receive either 20 weeks of CBT-E or IPT-ED (see Chapter 1 for a detailed description of these treatments). Following treatment, patients entered a closed follow-up period in which no other treatment was received. Due to the nature of treatment, it was not possible to blind therapists or patients to treatment allocation. Assessments took place at baseline (prior to starting treatment), at the end of treatment, and at 20 weeks, 40 weeks and 60 weeks follow-up. All assessments were undertaken by trained, independent assessors who had not been involved with treatment, who were blinded to treatment allocation.

3.2.1.3.1. Therapists

Patients were allocated to one of three therapists (2 clinical psychologists and 1 psychiatric nurse practitioner) who delivered both treatments. All therapists were specialists in eating disorders. One therapist had prior experience using CBT-E and another of interpersonal psychotherapy (IPT). All therapists received six months

training in CBT-E and IPT-ED prior to starting the trial and were supervised on a weekly basis throughout the trial for both treatments.

3.2.1.3.2. Treatment Fidelity

All treatment sessions were recorded. A random selection of session recordings were audited each week to ensure ongoing treatment fidelity. On completion of the trial treatment fidelity was by assessed four independent raters[3] using adapted versions of two adherence scales used in a previous study comparing cognitive behaviour therapy (CBT) and IPT for bulimia nervosa (Loeb et al., 2005). One hundred and twenty treatment sessions were randomly selected, stratified by early, middle and late stages of treatment. Fidelity was scored on a scale of 1 to 7, with higher scores indicating greater treatment fidelity. Fidelity ratings indicated that both treatments were well implemented, with CBT-E scoring a mean of 6.41 and IPT a mean of 6.15 (personal communication, R. Jones, 13th May 2014).

3.2.1.4. Measures

3.2.1.4.1. Eating Disorder Examination 16.0 (EDE 16.0)

The EDE is a validated and widely used clinical interview that assesses eating disorder psychopathology and symptomatology (Fairburn, Cooper, & O'Connor, 2008). It produces a global score denoting general psychopathology, which can be broken down into four subscales that assess 'dietary restraint', 'eating concern', 'weight concern' and 'shape concern', with different eating disorder features scored on a scale of 0 to 6 and an average score is calculated. Higher scores indicate greater psychopathology

[3] Including the author of this thesis

based on how frequent the eating disorder feature was present during the last 28 days. An operational definition of remission that has been commonly used in the literature (e.g. Agras et al., 2000) is a score of 1.74 on the EDE, which is equivalent to one standard deviation above the community mean. Scores equal to or greater than 1.74 indicate the presence of clinically severe eating disorder psychopathology, scores below this cut-off are indicative of remission (Beglin, 1990).

3.2.1.4.2. Beck Depression Inventory II (BDI-II)

The BDI-II is a validated and widely used self-report questionnaire that assesses depression (Beck, Ward, Mendelson, Mock, & Erbaugh, 1961). It has 21 items, scored on a scale of 0-4, and gives a cumulative score out of 63. Higher scores indicate greater levels of depression. In the original scale, scores can be interpreted as follows; 0-13 minimal depression, 14-19 mild depression, 20-28 moderate depression, 29-63 severe depression.

3.2.1.4.3. Rosenberg Self-Esteem Scale (RSES)

The RSES is a validated and widely used 10 item self-report measure that assesses general self-esteem (Rosenberg, 1965). Items are rated on a scale of 0-3, giving a cumulative score of 0-30. Higher scores indicate greater self-esteem.

3.2.2. Power

The study was powered to detect a mean change of 0.45 points on the EDE, and to detect a difference in remission of 25% between treatments at the end of treatment based on previous research (Agras et al., 2000) in the area (Fairburn et al., 2015). This gave a total sample size of 130 patients, with 65 patients per treatment arm.

3.2.3. Statistical Analysis and Results of Main Trial

Results were analysed using multilevel modelling (with baseline constrained to be equal in the two treatments) using Stata version 13 (StataCorp, 2013). Overall, 22% of patients withdrew or dropped-out during treatment, these included 17 patients with BN, and 12 with EDNOS. Eating disorder psychopathology decreased in both treatments; however, change was greater in CBT-E compared to IPT-ED. By the end of treatment a greater proportion of patients in CBT-E met criteria for remission (based on the EDE clinical cut-off) compared to IPT-ED (66% vs. 33%). This pattern was also observed at 60-week follow-up (69% vs. 49%). For more detailed results please see full text article in appendix 1.

3.3. Mediation Studies

3.3.1. Research Design and Procedure

3.3.1.1. Design

Mediation studies were embedded within the main RCT. These studies were designed to assess purported mediators of both treatments using the other treatment as a comparator (Murphy et al., 2009). In CBT-E, the studies concerned purported mediators of specific procedures within this multi-component treatment (Murphy et al., 2009). In IPT-ED, the entire treatment was treated as a complete unit due to its singular focus on addressing interpersonal problems. See Table 7 for a comparison of treatment procedures in each treatment and their timings.

This thesis examines three of the core procedures in CBT-E; 'regular eating', 'weekly weighing' and 'shape checking', and what is arguably the key purported mediator of

IPT-ED; interpersonal problem severity. These studies will be described in more detail in later chapters.

3.3.1.2. Procedure

Purported mediators and outcomes of CBT-E were assessed on a weekly basis, while purported mediators of IPT-ED were assessed over a longer time period; at the start of treatment (week 3), just before the end of treatment (week 19) and at the 20-week and 40-week follow-ups (see Table 8).

After each treatment session, therapists documented what had happened and which procedures had been used. The timing of data collection was based on what was known about the time course of action of cognitive behaviour therapy (CBT) and interpersonal psychotherapy (IPT) based on previous research (Murphy et al., 2009). Namely, that change in eating disorder psychopathology was slower in IPT than in CBT when treating bulimia nervosa (Agras et al., 2000; Fairburn, Jones, et al., 1993), a trend that was also observed in the main trial (Fairburn et al., 2015). The exact procedure varied in each study depending on the mediator, and will be documented in more detail in later chapters.

Table 7: Treatment timings of enhanced cognitive behaviour therapy (CBT-E) and interpersonal psychotherapy for eating disorders (IPT-ED)

Treatment Details	Treatment Procedures	Timing of Interventions during Treatment																									
WEEK		1		2		3		4		5	6	7	8	9	10	11	12	13	14		16		18		20
SESSION		0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17		18		19		20
CBT-E TREATMENT STAGES		Stage 1								Stage 2		Stage 3								Stage 4					
CBT-E TREATMENT COMPONENTS	Regular eating			X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X						
	Weekly weighing		X		X		X		X	X	X	X	X	X	X	X	X	X	X						
	Shape Concern and shape checking									X	X	X	X	X	X	X	X	X	X						
	Dieting and dietary rules									X	X	X	X	X	X	X	X	X	X						
	Events, moods and eating									X	X	X	X	X	X	X	X	X	X						
	Relapse prevention																				X		X		X
IPT-ED TREATMENT PHASES		Phase 1						Phase 2												Phase 3					
IPT-ED	Identifying interpersonal problems	X	X	X	X	X	X																		
	Addressing interpersonal problems							X	X	X	X	X	X	X	X	X	X	X	X		X		X		X
	Relapse prevention																				X		X		X

Table 8: Timing of Measures

Details	Measures	Pre-Treatment	Timing of Assessments During Treatment																							Follow-Up	
WEEK			1		2		3		4		5	6	7	8	9	10	11	12	13	14		16		18		20	20	40
SESSION			0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17		18		19		20		
ASSESSMENTS	EDE	X																								X	X	X
	IPI						X																	X			X	X
	RERS		X		X		X		X		X																	
	BERS		X		X		X		X		X																	
	EQ		X		X		X		X		X	X	X	X	X	X	X	X	X	X		X		X		X		
	IQ		X		X		X		X		X	X	X	X	X	X	X	X	X	X		X		X		X		

Notes. IPI=interpersonal interview, EDE=eating disorder examination, RERS=regular eating rating scale, BERS= binge eating rating scale, EQ=eating questionnaire, IQ=interpersonal questionnaire

3.3.1.3. Measures

Basic descriptions of the measures are provided below, however, additional information regarding timing of measurement and measure reliability will be given in the relevant chapters.

3.3.1.3.1. Regular Eating Rating Scale (RERS)

The RERS is a rating scale that assesses adherence to a regular eating pattern (based on patients' daily self-monitoring records) of 3 meals plus 2/3 snacks daily (with no greater than 4 hours between each meal/snack) (see Appendix 3 and Appendix 4). This is excluding binge eating episodes, which are counted as periods of irregular eating and do not contribute as meals/snacks. This scale was developed for use in the current study. Adherence is rated on a scale of 0 to 6, with higher scores indicating greater regularity of eating.

3.3.1.3.2. Binge Eating Rating Scale (BERS)

The BERS is a rating scale that assesses frequency of binge eating (based on patients' daily self-monitoring records) (see Appendix 4 and Appendix 5). Binge eating (BE) was defined as eating a substantially larger amount of food than most people would eat during a 2 hour period. This scale was developed for use in the current study. The rating scale assesses number of binge episodes per day, and then calculates a daily average.

3.3.1.3.3. Eating Questionnaire (EQ)

The EQ is a 31 item self-report questionnaire assessing eating disorder psychopathology and symptomatology (see Appendix 6). This scale was developed for

use in this study and was adapted from the questionnaire version of the EDE described above (Fairburn & Beglin, 2008). Phrasing and content were adjusted so that it better captured purported mediators and outcomes, and additional items were also added to assess purported mediators of CBT-E that were not already assessed in the original EDEQ. Items are rated on a scale 0-7, representing 0 to 7 days. Scores are interpreted similarly to the EDE, with higher scores indicating greater psychopathology.

3.3.1.3.4. Interpersonal Questionnaire (IQ)

The IQ is a 21 item self-report questionnaire that includes a selection of items related to interpersonal difficulties, self-efficacy/self-confidence, and low mood (see Appendix 7). Items are rated on a Likert scale ranging from 0 to 6, where 0 represents 'not at all' and 6 'very much so'. This questionnaire was developed for the present study. It was used as a design feature so that patients in both conditions were being asked about both eating disorder psychopathology and interpersonal functioning on a weekly basis (which otherwise would not have been the case). These data were not analysed as the questionnaire did not derive from a validated measure.

3.3.1.3.5. Interpersonal Interview (IPI)

The IPI is a semi-structured interview that assesses interpersonal problems and interpersonal self-efficacy over the preceding four weeks (see Appendix 8). The IPI was developed for the current study. Patients identify up to three interpersonal problems prior to interview using the 'Interpersonal Problem Form' (see Appendix 9), with one interpersonal problem being identified as the key interpersonal problem. These are then categorised into different problem areas. The interview then focuses on these problem areas. The IPI gives three interviewer-rated scores; interpersonal

problem severity, problem-specific self-efficacy (for addressing the identified problem), and general interpersonal self-efficacy. Higher scores indicate greater interpersonal problem severity, and greater self-efficacy.

3.3.1.3.6. Intervention Record Sheet (IRS)

The IRS is a template for recording what happened in each treatment session. An IRS template was created for both CBT-E (see Appendix 10) and IPT-ED (see Appendix 11), which recorded stage/phase of treatment, session number, treatment week, aims, content and actions. In addition, the CBT-E IRS also included a checklist to record which interventions were used in the session in order to track personalisation of treatment in stage 3 of CBT-E.

3.3.1.4. Power

In the main trial, power was calculated on the basis of the sample size needed to detect a mean difference between CBT-E and IPT-ED in remission and significant change in scores over time (see 3.2.2). It was not powered on the basis of sample size needed to detect a mediated effect. The mediation studies are therefore likely to be underpowered and consequently their analysis was conducted on an exploratory basis. This issue will be discussed later in chapter 4 and chapter 9.

3.4. Aims of Thesis

The aim of the research described in this thesis was to investigate whether the hypothesised mechanisms of action (i.e. purported mediators) of CBT-E and IPT-ED mediated the effect of treatment on the outcome. Such research requires complex statistical modelling.

Specific objectives were to;

- Develop an appropriate analytic strategy for statistical mediation in the context of an RCT comparing two active psychological treatments using longitudinal data
- Develop an approach for assessing statistical mediation in the context of treatment personalisation within the analytic strategy developed
- Assess purported mediators of CBT-E and IPT-ED using the analytic strategy developed

The following chapter examines the complexities and practicalities of undertaking such an analysis and describes the over-arching analytic strategy developed.

Chapter 4. Statistical Methods for Investigating Purported Mechanisms of Action of Treatment

In the context of treatment, statistical mediation analysis can be used to consider how a treatment, its purported mediator, and the outcome are related to each other in a statistical sense (Baron & Kenny, 1986). There are currently two main schools of thought regarding statistical mediation analysis (Dunn et al., 2015; Preacher, 2015). The first approach traditionally involves an examination of the observable relationships between variables, using statistical methods such as multiple regression and structural equation modelling. The second approach, applies more modern conceptualisations of causal inference, and considers the problem of statistical mediation as an estimation of the causal effects of treatment.

4.1. Traditional Approaches to Statistical Mediation Analysis

4.1.1. Causal Steps Approach

4.1.1.1. Multiple Regression

The causal steps approach of Baron and Kenny (1986) is currently the most widely cited approach for undertaking statistical mediation analysis, at least in the field of psychological research (Kline, 2015; MacKinnon et al., 2007). This approach traditionally involves assessing for statistical mediation using three standard regression models. These regression models are used to answer questions regarding the

presence or absence of the relationships that would be expected to be observed under conditions of statistical mediation. There are four criteria that need to be met.

4.1.1.1.1. Criterion 1: Treatment Predicts the Outcome

The first regression model assesses whether treatment is related to the outcome (see Figure 8). In this regression treatment represents the 'total effect' of treatment on the outcome, including those transmitted through the purported mediator as well as any other processes. Treatment (X) should significantly affect the outcome (Y), as denoted by coefficient (*c*) in the model, with $p<.05$.

$$\underset{\text{[Outcome}}{Y} = \underset{\text{Intercept}}{i_1} + \underbrace{\underset{\text{Treatment}}{cX}}_{\text{Overall effect of treatment on the outcome ('total effect')}} + \underset{\text{Error]}}{e_1}$$

Figure 8: Regression model 1 in the causal steps approach

4.1.1.1.2. Criterion 2: Treatment Predicts the Purported Mediator

The second regression model assesses whether treatment is related to the purported mediator (see Figure 9). Treatment (X) should significantly affect the purported mediator (M), as denoted by coefficient (*a*) in the model, with $p<.05$.

$$\underset{\text{[Purported Mediator}}{M} = \underset{\text{Intercept}}{i_2} + \underbrace{\underset{\text{Treatment}}{aX}}_{\text{Effect of the purported mediator on the outcome}} + \underset{\text{Error]}}{e_2}$$

Figure 9: Regression model 2 in the causal steps approach

4.1.1.1.3. Criterion 3: The Purported Mediator Predicts the Outcome

The third regression model includes both treatment and the purported mediator in the

same equation (see Figure 10). This regression assesses whether the relationship between treatment and the outcome might be explained by the purported mediator, firstly by assessing whether there is a relationship between the purported mediator and the outcome. For criterion 3, the purported mediator (M) should significantly affect the outcome (Y), as denoted by coefficient (b) in the model, with $p<.05$.

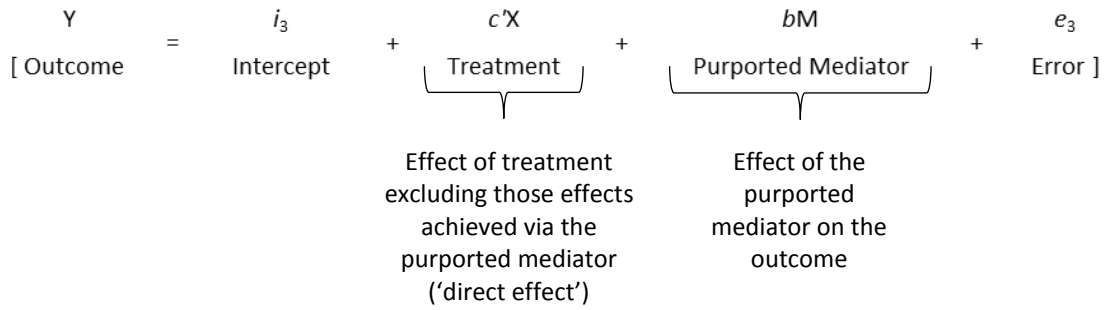


Figure 10: Regression model 3 in the causal steps approach

4.1.1.1.4. Criterion 4: The Impact of Treatment on the Outcome is Reduced When the Purported Mediator is taken into Account

The second purpose of the third regression model (see Figure 10) is to assess the impact of treatment on the outcome whilst controlling for the effect of the purported mediator. This regression therefore examines the direct effect of treatment on the outcome excluding any effect of treatment transmitted via the purported mediator. For criterion 4, the effect of treatment (X) on the outcome (Y) should change when the purported mediator (M) is accounted for, as denoted by (c') in the model. For full mediation, (c') is reduced to the point that the effect of treatment on the outcome is no longer significant, with $p>.05$, suggesting that the effect of treatment on the outcome is fully explained by the purported mediator. When this is the case a coefficient of zero provides the strongest support for full mediation. Under conditions

of partial mediation, (c') is reduced, but the effect of treatment remains significant, with $p<.05$, suggesting that the purported mediator only partly accounts for the effect of treatment on the outcome. The latter scenario is likely when considering complex psychological treatments as it is unlikely that they work solely via one process. According to Baron and Kenny (1986), all three regression models must show the expected effects for mediation to be supported.

4.1.1.1.5. Testing for Moderators

Baron and Kenny (Baron & Kenny, 1986) also provide guidance for testing potential moderators of treatment's effect on the outcome. This is done by testing an interactive effect between treatment and the potential moderator (see Figure 11). For the potential moderator (Z) to be accepted as a moderator the effect of treatment (X) on the outcome (Y), the interaction between these variables (XZ) should significantly affect the outcome (Y), as denoted by coefficient (c_3) in the model, with $p<.05$.

$$\underset{\text{Outcome}}{Y} = \underbrace{i_1}_{[\text{ Intercept}} + \underbrace{c_1X}_{\text{Treatment}} + \underbrace{c_2Z}_{\text{Moderator}} + \underbrace{c_3XZ}_{\text{Treatment X Moderator}} + \underbrace{e_1}_{\text{Error }]}$$

Effect of treatment in absence of moderator

Effect of moderator in absence of treatment

Effect of moderator on treatment A vs. B

Figure 11: Regression model for testing treatment moderators

4.1.1.2. Extension to Multilevel Modelling

The extent to which standard regression models can incorporate longitudinal data is limited (Krull & MacKinnon, 1999, 2001). This is because the inclusion of repeated measures data within the same regression model violates the assumption of independence of errors (Krull & MacKinnon, 1999). However, the causal steps

approach of Baron and Kenny (1986) can be extended to multilevel modelling (MLM) to take into account the nested structure of repeated measurements within the same patient (Krull & MacKinnon, 1999, 2001). In MLM, model equations specify variables at the different levels of the data's nested structure e.g. at the between patient or within patient level (see Figure 12).

Whilst MLM is an extension of standard regression, these analyses differ in important ways (Tabachnick & Fidell, 2005). MLM analyses are more complex, taking into account not just the variance, but also the covariance between error terms. Model estimation is iterative and is typically undertaken using a maximum likelihood algorithm to fit the model to the data. In repeated measures designs, time is explicitly modelled. This offers greater flexibility in modelling trends over time (e.g. linear, quadratic). It is also possible to model additional time trends within the same model, for example, denoting the occurrence of specific events during treatment. These are known as piecewise trends (McCoach & Kaniskan, 2010; Singer & Willett, 2003). Variables can be entered into the model as fixed and/or random. When a variable is specified as fixed (default), its effect is set to be the same across all participants. However, when a variable is specified as random it is allowed to vary across all participants. An example of when the latter might apply is the case of time. As a fixed variable, time would assess a steady rate of growth at each time point across all participants. As a random variable, this would assess whether rate of growth differed across participants. Their ability to incorporate multiple repeated measurements also means that they are much more useful than multiple regression for studying purported mediators.

Multilevel Model 1: Treatment predicts the outcome

Patients (Level 2, $_j$)	Y_{ij} [Outcome	=	i_{0j} Between patients Intercept	+	cX_j Treatment	+	u_{ij} Between Patients Error]	
Within Patient (Level 1, $_i$)	i_{0j} [Between Patients Intercept	=	i_{ij} Within Patient Intercept	+	e_{ij} Within Patient Error]			

Multilevel Model 2: Treatment predicts the purported mediator

Patients (Level 2, $_j$)	M_{0j} [Purported Mediator	=	i_{0j} Between patients Intercept	+	aX_j Treatment	+	u_{ij} Between Patients Error]	
Within Patient (Level 1, $_i$)	i_{0j} [Between Patients Intercept	=	i_{ij} Within Patient Intercept	+	e_{ij} Within Patient Error]			

Multilevel Model 3: Treatment predicts the purported mediator

Patients (Level 2, $_j$)	Y_{0j} [Outcome	=	i_{0j} Between patients Intercept	+	$c'X_j$ Treatment	+	u_{ij} Between Patients Error]	
Within Patient (Level 1, $_i$)	i_{0j} [Between Patients Intercept	=	i_{ij} Within Patient Intercept	+	bM_{ij} Purported Mediator	+	e_{ij} Within Patient Error]	

Figure 12: Multilevel Modelling Equations for a 2-1-1 mediation model in which treatment is randomised and both the purported mediator and the outcome are repeated measures.

Due to the hierarchical nature of MLM the variables implicated in the mediation model may occur at different levels of the data structure (Bauer, Preacher, & Gil, 2006; Krull & MacKinnon, 1999). For example, in Figure 13, treatment allocation occurs at level 2 in the model because treatment is randomised at the patient level. In contrast, time occurs at level 1 as it describes the within-patient experience. Both the purported mediator and the outcome also occur at level 1 as they are repeated measures. This is an example of a 2-1-1 multilevel mediation model.

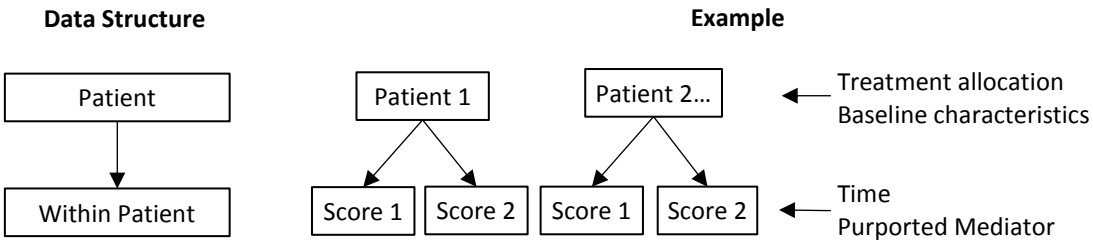


Figure 13: Example Data Structure of a 2-1-1 Multilevel Mediation Model

Whilst a 2-1-1 structure is likely to be the case in most treatment RCTs, in the context of treatment personalisation a different model applies. When treatment procedures are implemented at different times for different patients (as in enhanced cognitive behaviour therapy), treatment is effectively being manipulated at the within-patient level, not just at the patient level. Such a scenario can be conceptualised as treatment personalisation at the within-patient level being nested within the between-patient variable of treatment allocation. When treatment, the purported mediator and the outcome are measured at the within-patient level this is a 1-1-1 multilevel mediation model. For models in which both the purported mediator and the outcome occur at level 1 (i.e. 2-1-1 and 1-1-1 models), the purported mediator is entered as a 'time-varying' covariate. When this is the case, it is possible to assess the impact of the purported mediator on the outcome based on their contemporaneous relationship

(i.e. their relationship within the same time point) and their lagged relationship (i.e. the relationship prior level of the purported mediator has on subsequent outcome). The latter enable a more thorough test of the temporal relationships between variables, and therefore a more thorough test of the temporal relationships implied in statistical mediation. An alternative model would be a model in which the purported mediator occurred at the patient level, rather than the within-patient. This is an example of a 2-2-1 multilevel mediation model.

Whilst 2-1-1, 1-1-1 and 2-2-1 models might examine the same purported mediator, they represent very different conceptualisations of the purported mediator. The 2-1-1 and 1-1-1 models assess the dynamic relationship between the purported mediator and the outcome, looking at how fluctuations in the purported mediator relate to the outcome. In contrast, the 2-2-1 model, examines how the average level of the purported mediator during treatment relates to the outcome. Such a model might be of interest when fluctuations during treatment are hypothesised to be less important than the overall level of the purported mediator across the entirety of treatment. Due to the hierarchical structure of the data, even when treatment is manipulated at the within-patient level, it is not possible to have the purported mediator specified at a higher level of the model (Zhang, Zyphur, & Preacher, 2009). The treatment effect transmitted via the purported mediator can only travel downwards through the different levels of the model.

4.1.1.3. Updated Causal Steps Approach

Whilst the causal steps approach of Baron and Kenny is arguably the leading approach in statistical mediation currently, the approach described in their landmark article of

1986 (Baron & Kenny, 1986) has been criticised by a number of authors. This has resulted in three major revisions to the statistical methods.

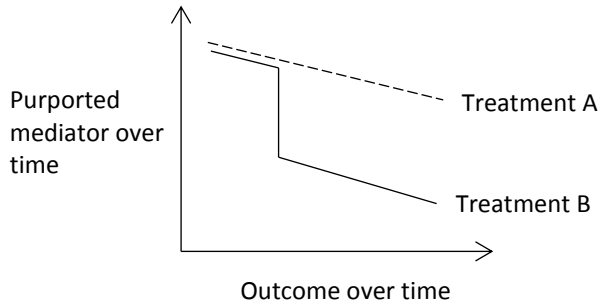
4.1.1.3.1. Revision to Criterion 1: Treatment May Not Predict Outcome when Competing Mediators Exist

Whilst the Baron and Kenny approach requires a significant relationship between treatment and the outcome in regression model 1, it is possible that genuine mediators may not always fulfil criteria 1 (MacKinnon, Krull, & Lockwood, 2000; Shrout & Bolger, 2002). For example, if the treatment affects multiple mediators that have oppositional effects (i.e. one increases the outcome, the other decreases it), these competing mediating effects will cancel out each other's effect on the outcome giving an overall null effect. By failing to take into account such effects in the causal steps process, genuine mediators may be incorrectly rejected. It has therefore been advised that statistical mediation not be rejected purely on the basis of a non-significant relationship between treatment and the outcome.

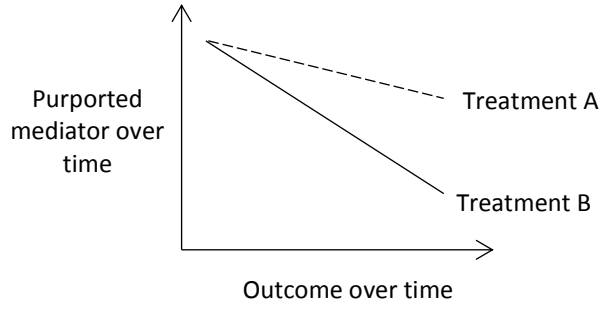
4.1.1.3.2. Revision to Criterion 3: Assess for Interactions between Treatment and the Purported Mediator

The Baron and Kenny approach has also been criticised for ignoring the possibility of an interaction between treatment and the purported mediator (Kraemer et al., 2002). In the original approach the purported mediator is tested only as main effect. This assumes that the purported mediator enacts in similar ways across treatments (see graph A, Figure 14). However, this may not be the case (Kraemer et al., 2002). Treatment could enact by changing the overall level of the purported mediator in the treatment group, as a 'main effect' only (see graph A, Figure 14). Or it could change

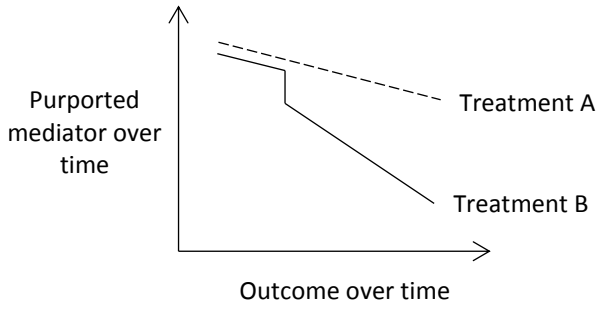
how the purported mediator affects the outcome as an 'interactive effect' so that change in the purported mediator affects the outcome faster (see graph B, Figure 14). Or it could do both (see graph C, Figure 14).



A. Difference in the level of the purported mediator between treatments



B. Difference in how the purported mediator relates to outcome in the two treatments



C. Difference in the level of the purported mediator and how it relates to outcome in the two treatments

Figure 14: Illustration of how treatment's effect on the purported mediator and consequently the outcome might vary between treatments

Excluding an interaction effect when there is one increases bias in the model and reduces statistical power (Kraemer, 2010). It has therefore been suggested that an interaction between treatment and the purported mediator be routinely tested as part of regression model three (see Figure 15). This is essentially the same as testing treatment as a moderator of the effect of the purported mediator on the outcome. For treatment (X) to moderate the effect of (M) on the outcome, the interaction between these variables (XM) should significantly affect the outcome (Y), as denoted by a coefficient (c_2) in the model, with $p<.05$. Similarly, interaction effects can be fitted in MLM.

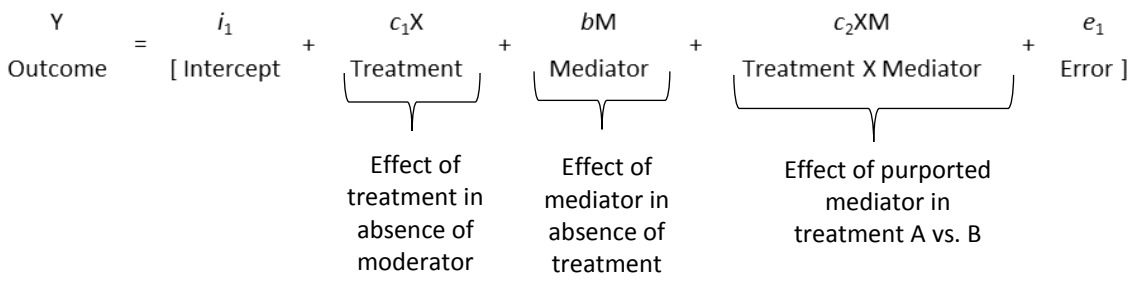
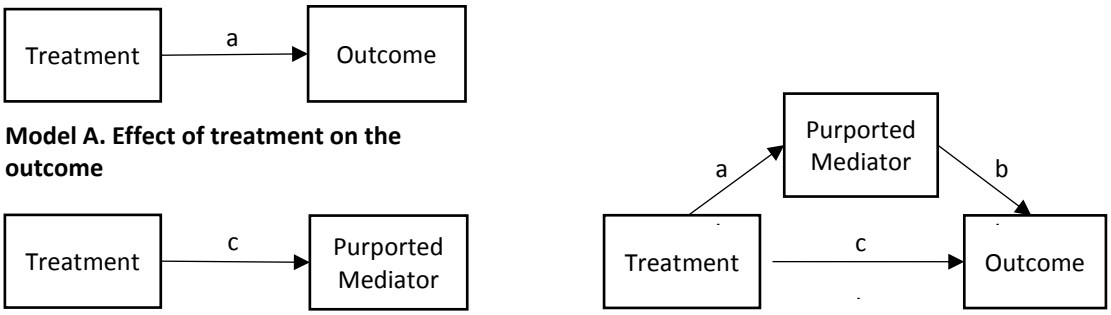


Figure 15: Modified regression model 3 in updated causal steps approach testing for an interactive effect between treatment and the purported mediator

4.1.1.3.3. Revision to Criterion 3: Estimating a Mediated Effect

The original Baron and Kenny causal steps approach to statistical mediation determines mediation based on the statistical significance of the model coefficients of the three regression analyses outlined previously. However, *p* values have limited value for understanding the clinical significance of a mediated effect in terms of how effective it is for a treatment to target the purported mediator (Kraemer et al., 2002). A more useful measure is the 'mediated' or 'indirect effect', which attempts to quantify the effect that treatment has on the outcome via the purported mediator. Several methods for calculating the mediated effect have been proposed that can be

used when both the purported mediator and the outcome are continuous variables (MacKinnon, Lockwood, Hoffman, West, & Sheets, 2002). The most reliable of these methods include the difference in coefficients method and the product of coefficients method. The difference in coefficients method can be calculated as the difference between the total effect of treatment on the outcome and its direct effect (paths c-c' in Figure 16). The product of coefficients method is calculated as the effect that treatment has on the purported mediator multiplied by the effect that the purported mediator has on the outcome (paths a'b' in Figure 16). The difference between these two methods is subtle, and they give equivalent results when the same patients are used in all steps of the model (i.e. variables are not affected differently in terms of missing data). As such the product of coefficients might be considered preferable as it directly computes the effect of treatment via the purported mediator, as opposed to making an indirect estimate based on changes in the effect of treatment on the outcome when controlling for the purported mediator.



Model A. Effect of treatment on the outcome

Model B. Effect of treatment on the purported mediator

Model C. Decomposed model. Direct and indirect effect of treatment on the outcome

Figure 16: Decomposition of the coefficients (represented as paths) needed to calculate the total, direct and indirect effects of treatment

There exist a number of methods for testing the statistical significance of the mediated effect. The most well-known is the Sobel test (Sobel, 1982), a method advocated by Baron and Kenny (1986). These tests have also been criticised for their reliance on *p*

values. The use of confidence intervals to quantify the size and spread of the mediated effect has been recommended, with significance based on whether the confidence interval crosses 0 (Kraemer, 2008). If 0 is within the range then the mediated effect is not significant.

In MLM, the mediated effect can be calculated using similar methods to those of regression (MacKinnon, 2008). However, quantifying the mediated effect in such models is more complicated as such calculations need to take into account how treatment and the purported mediator enact across the different levels of the data structure (Zhang et al., 2009). The specification of random effects also adds to the complexity of the model and its subsequent interpretation (Kenny, Korchmaros, & Bolger, 2003).

4.1.1.3.4. Additional Assumptions in Statistical Mediation Analysis

Statistical mediation using the causal steps approach involves the usual assumptions of standard regression models i.e. linearity, homoscedasticity, independence of errors, no missing data, no or low measurement error, and no feedback loops between the outcome and other variables in the model (i.e. treatment, the purported mediator). The handling of missing data (e.g. listwise deletion, multiple imputation) and reliability of measures therefore needs to be taken into account. Measurement error is particularly problematic in the purported mediated as it simultaneously leads to an underestimation of the effect treatment on the purported mediator and the effect of the purported mediator on the outcome, and consequently leads to an over-estimation of the direct effect of treatment (Baron & Kenny, 1986; Cole & Maxwell, 2003; Judd & Kenny, 1981). Collinearity between the purported mediator and the

outcome also needs to be considered, as, whilst indicative of a strong relationship, it increases estimate biases, resulting in reduced power to detect a significant effect (Judd & Kenny, 1981).

Statistical mediation is particularly vulnerable to confounding due to the multiple relationships being modelled. However, this risk is reduced when treatment and the purported mediator are randomised, as in a double randomisation randomised controlled trial (RCT). Assuming randomisation has adequately balanced all potential (causally) confounding variables across conditions then it can be assumed that any change in the outcome or the purported mediator was due to treatment. Similar causal inferences can be made about the effect of the purported mediator on the outcome when the purported mediator has been randomised. However, perfect randomisation often fails (Rubin, 1974). Furthermore, when the purported mediator has not been randomised (as is more commonly the case), the extent to which change in the outcome can be attributed to the purported mediator is less sure. Explicit consideration of potential confounders and how they are controlled for is therefore an important part of statistical mediation analysis. Adjusting for potential confounders in the analysis is one potential solution to this problem. However, this can become complicated when confounders represent chains of mediators that affect each other during the course of treatment. This affects the estimation of the direct effect of treatment on the outcome, and therefore calculation of the mediated effect (Daniel, De Stavola, Cousens, & Vansteelandt, 2015; Keele, 2015). An important tool in the identification of potential confounders is the directed acyclic graph (Greenland, Pearl, & Robins, 1999), which is a visual representation of the hypothesised relationships

between treatment, the purported mediator and the outcome, as well as additional variables that may confound these relationships (e.g. 'common causes', additional mediators). This includes consideration of relevant non-specific predictors and potential moderators. The inclusion of relevant non-specific predictors and moderators also reduces the amount of unexplained variance in the data thus improving the reliability of estimates and increasing power (Dunn et al., 2015; Fairchild & MacKinnon, 2009; Fritz, Cox, & Mackinnon, 2013).

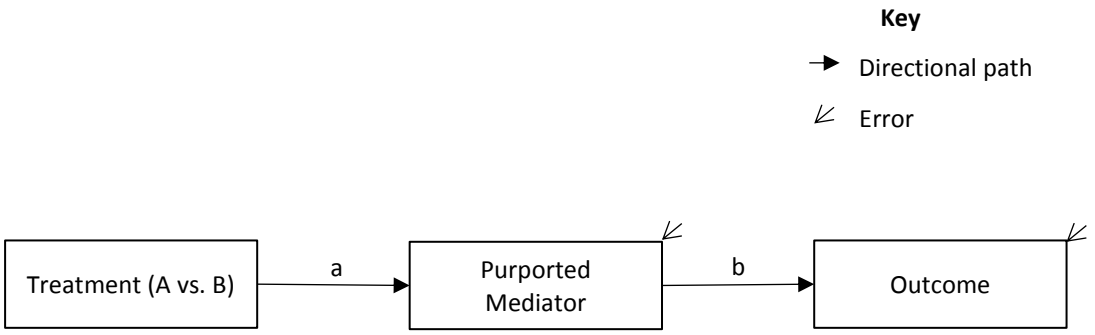
4.1.2. Structural Equation Modelling

One of the limitations of the causal steps approach is that the relationships between treatment, the purported mediator and the outcome are tested separately in a series of univariate models, rather than being assessed in one model simultaneously. This can result in inaccurate estimates (Judd & Kenny, 1981). For this reason structural equation modelling (SEM) has been recommended over the causal steps approach due to its ability to model the sequential relationships between treatment, the purported mediator and the outcome simultaneously in one multivariate model (Iacobucci, Saldanha, & Deng, 2007; Judd & Kenny, 1981). Whilst SEM makes a greater number of assumptions regarding the causal relationships between variables, they produce more accurate estimates of the mediated effect under conditions of partial mediation (James, Mulaik, & Brett, 2006).

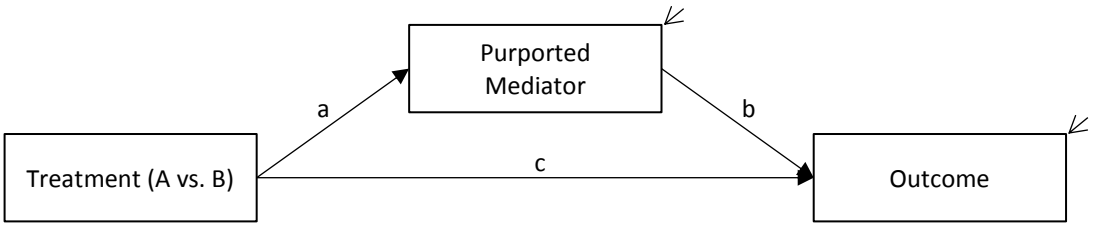
SEM encompasses a variety of related regression-based techniques, and therefore has the same assumptions as multiple regression i.e. linearity, normality, homoscedasticity, and independence. Although based on regression, these techniques are considerably more complex, as they take into account the covariance structure of

the data (i.e. the interrelationships between variables). Whilst each technique has slightly different processes and assumptions, there are five key steps that are common to all of them. These are model specification, model identification, model estimation, model testing and, if appropriate, model modification (Kline, 2011).

Model specification involves specifying the variables in the model and how they relate to each other. This is done by specifying 'paths' between variables that denote the direction of their relationship. The simplest SEM technique is 'path analysis', which directly models relationships between observed variables. SEM is typically depicted using a visual representation of the model (see Figure 17). Paths include directional relationships (typically depicted using a single headed arrow in the direction of the hypothesised relationship) and covariances (typically depicting using double headed arrows).



Model A. Full mediation path analysis



Model B. Partial mediation path analysis model

Figure 17: Path analysis models for full and partial mediation

Both the presence and absence of paths between variables represent assumptions being made in the model. The absence of a path implies no relationship between variables. A weaker assumption than no path is specifying a directional path between variables, although this obviously implies a causal relationship. The weakest assumption is specifying covariance between variables, which implies that a relationship may exist between these variables, but the causal direction is unknown. Specification is carried out a priori; this means that full or partial mediation must be specified from the start, unlike the causal steps approach, which assesses full and partial mediation the same way (James et al., 2006). A distinction is made between 'exogenous' variables that are not predicted by any other variable in the model (e.g. treatment in Models A and B in Figure 17), and endogenous variables that are predicted by other variables in the model (e.g. the purported mediator and the outcome in Models A and B in Figure 17). All exogenous variables are assumed to co-vary. All endogenous variables have an associated error term (see Figure 17).

Although statisitical mediation in SEM is affected by the same assumptions described earlier (i.e. measurement error, confounding), one advantage of SEM is that it is able to model measurement error explicitly. This is done using 'latent' variables to represent the underlying construct/s (e.g. depression, eating disorder psychopathology). They are estimated based on observed variables (e.g. scores on individual questionnaire items) that act as 'indicators' of the underlying construct (Iacobucci et al., 2007; James et al., 2006; Judd & Kenny, 1981). This technique typically requires at least three items to avoid imposing strong assumptions on the model, however (Little, 2013). Alternatively, items can be combined (or 'parcelled')

together into a single score, normally an average. Parcelled variables have stronger psychometric properties (e.g. reliability), but do not model measurement error directly. They are suitable when items represent a univariate measure i.e. a single construct only (Little, Cunningham, Shahar, & Widaman, 2002; Little, Rhemtulla, Gibson, & Schoemann, 2013).

Model identification involves assessing whether there are enough degrees of freedom in the model for the model to be estimated. For the model to be estimated, the number of parameters to be estimated (i.e. paths, covariances, variances of exogenous variables and error terms of endogenous variables) must not exceed the number of observations in the model (calculated as v(v+1)/2 where v is the number of observed variables). When the number of parameters is fewer than the number of observations the model is considered to be 'over-identified', meaning that it has at least one degree of freedom. For models in which the number of parameters equals the number of observations, model estimation is restricted, and degrees of freedom equal zero. Model A in Figure 17 represents an over-identified model, as there are five parameters (paths a, b, the variance of treatment, and the error terms of the purported mediator and the outcome) and three observed variables (so number of observations is 3(3+1)/2 = 6). It therefore has one degree of freedom. In contrast, whilst Model B has the same number of observations, it specifies one additional parameter (path c') and is therefore only just-identified.

Model estimation is typically done using maximum likelihood estimation techniques in which the model is fit against the data using an iterative process. Variances, covariances, path coefficients and error terms are calculated simultaneously until the

best fitting solution to the model is found. Path coefficients with a *p* value less than .05 (or confidence intervals that do not cross zero) suggest the relationship is making a significant contribution to the model. Mediated effects can be calculated using the product of coefficients method by multiplying the relevant paths between treatment and the purported mediator, and the purported mediator and the outcome.

Model testing involves assessing goodness-of-fit indices which compare the specified model to what was actually observed in the data. There is some controversy surrounding the utility of these indices. However, current recommended practice is to report multiple fit indices as different fit indices have different strengths and limitations (Boomsma, 2000; Hu & Bentler, 1999; Kline, 2011; Rigdon, 1996). A summary of key fit indices and their strengths and weaknesses can be found in Table 9. Specificity of the model (i.e. the extent to which the specified model is the best fit for the data) can also be tested by assessing plausible alternative models (Kline, 2011). Where the alternative model is 'nested' within the original model, model fit can be compared using a chi-square test. Most SEM programmes produce 'modification indices' for how model fit could be improved by adding or removing paths based on the chi-square test. Model modification should be based on a theoretical rationale, rather than being solely guided by the modification indices produced by the programme. This is to avoid adding paths that make no theoretical sense (e.g. directional paths where end of treatment scores predict baseline).

Longitudinal data can be modelled in a variety of ways to examine both differences between individuals at different time points, as well as change over time (Cheong, MacKinnon, & Khoo, 2003; Cole & Maxwell, 2003; MacKinnon, 2008). Two of the more

commonly used models are autoregressive SEM and latent growth curve modelling. Whilst other models exist, these will not be discussed here.

Table 9: Summary of Four Common Fit Indices, their Cut-Offs and their Relative Strengths and Limitations

Fit Indices	Indicators of Good Fit	Strengths	Weaknesses
Chi-square test (Chi2)	Chi2, $p > .05$	Can be used to directly test fit between nested models.	Greater type 1 error (false fit) with non-normal data, when higher correlations among observed variables and measure reliability is low. Greater type 2 (false rejection) errors in large sample sizes (N > 4000).
Comparative Fit Index (CFI)	CFI > .95	Less affected by non-normal distribution & sample size (including small sample size). Sensitive to model misspecification (although these attributes are controversial).	Fit index is calculated by comparing the specified model to a 'baseline' model of no relationships (unrealistic model of comparison).
Root Mean Square Error of Approximation (RMSEA)	RMSEA < .08 *p*close < .05	Supposed to correct for model complexity. Sensitive to model misspecifications (although these attributes are controversial).	May be affected by non-normal distribution & sample size (being less accurate in small sample sizes).
Akaike Information Criterion (AIC)	Smaller AIC	Can be used to directly compare fit between non-nested models.	Value depends on the number of parameters specified in a model.

Notes. Table summarises commentaries from a variety of research sources (Boomsma, 2000; Fan & Sivo, 2005; Feinian, Curran, Bollen, Kirby, & Paxton, 2008; Hu & Bentler, 1999; Kenny & McCoach, 2003; Kline, 2011; Marsh, Hau, & Wen, 2004; Rigdon, 1996)

4.1.2.1. Autoregressive Structural Equation Models

Autoregressive SEM examines the relationships between different variables over time by modelling their relationships at different time points. Its name ('autoregressive') refers to the way in which the model handles repeated measures of the same variable, with prior measurements of the same variable predicting subsequent measurements (essentially regressing the variable on itself). This modelling technique therefore controls for the influence of prior levels of the same variable. Direct relationships are specified between proximal/adjacent time points only (see Figure 18). An implicit assumption in the model is that the influence of earlier time points diminishes as they become more distal, their influence being conveyed via the autoregressive path. Due to the fact that repeated measures are likely to have shared variance, it is recommended that the measure error terms at each time point be correlated to prevent over-estimation of the path coefficients (Cole & Maxwell, 2003; Marsh, 1993).

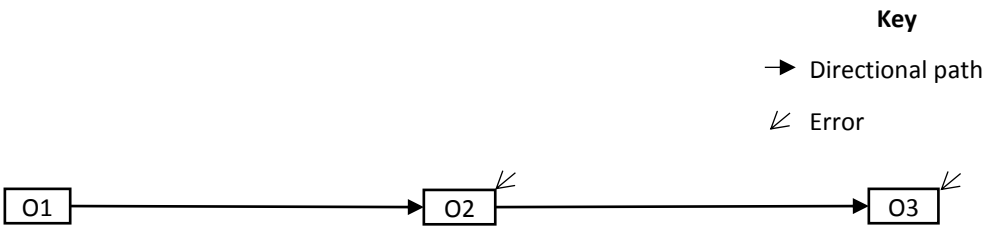


Figure 18: Example of an autoregressive model of Outcome O at time points 1-3

Autoregressive SEM is typically used in situations where the relationships have reached a certain level of constancy (Hertzog & Nesselroade, 1987). This applies to its mean ('stability'), its variance and covariance ('equilibrium') and its causal relationships with other variables (stationarity). This is in order to make interpretation meaningful where relationships are merely being observed. However, in situations where a variable has been experimentally manipulated (e.g. receipt of a treatment) it

can be a useful method for examining the impact of the manipulation, both immediately and at later time points. In these situations interpretation of the path coefficients on the autoregressive path has certain implications; a coefficient of close to 0 implies instability in the variable (because there is little relationship between the prior and subsequent time points), a coefficient of 1 implies stability. The stability of path coefficients will be affected by measurement error. Where this is high, path coefficients are likely to be lower due to a lack of precision of the measurements. In such scenarios, timing of intervals of data collection is crucial to modelling the impact of the experimental manipulation as this will affect magnitude of effect; if intervals are two wide then instability on the model may not be detected. Selection of intervals should be based on timings of theoretical interest (Cole & Maxwell, 2003). For example, treatment week.

4.1.2.1.1. Autoregressive Models for Statistical Mediation

In the context of statistical mediation, autoregressive paths are specified for treatment, the purported mediator and the outcome; except when change in treatment status is a one-off event that occurs at the same time for all patients. For example, such as when treatment allocation is randomised at baseline. Baseline variables (including treatment allocation) are treated as exogenous variables and are correlated at baseline. This also effectively controls for baseline relationships between treatment allocation, the purported mediator and the outcome. In theory, some degree of relationship should be observed between the purported mediator and the outcome at baseline, unless treatment is hypothesised to change the relationship between these two variables in some way. There may or may not be an association

between treatment and the purported mediator / outcome depending on whether treatment was allocated on the basis of existing characteristics (i.e. specific levels of the purported mediator / outcome) or by chance (in the case of randomisation).

Paths can be specified between treatment, the purported mediator and the outcome in a number of ways, depending on whether full or partial mediation is hypothesised, and their hypothesised relationships at each time point. Specifically, relationships between the purported mediator and the outcome can be modelled as longitudinal, contemporaneous, or cross-lagged (MacKinnon, 2008) (see Figure 19). In longitudinal mediation the purported mediator at an earlier time point predicts the outcome at a later time point. In contemporaneous mediation the purported mediator is associated with or directly related to the outcome at the same time point. For example, an association maybe specified when causal direction is unknown or to control for contemporaneous associations when modelling only longitudinal relationships. Contemporaneous paths may be specified in situations where change is occurring more rapidly than is captured by the longitudinal paths. Finally, cross-lagged mediation implies a reciprocal relationship between the purported mediator and the outcome, whereby prior level of the purported mediator predicts the outcome at the subsequent time point, and vice versa (prior levels of the outcome predict the purported mediator at the subsequent time point). Although it is possible to specify reciprocal relationships within the same time point, this complicates both the model estimation and calculation of indirect effects due to the potential for infinite feedback loops (Kline, 2011; MacKinnon, 2008).

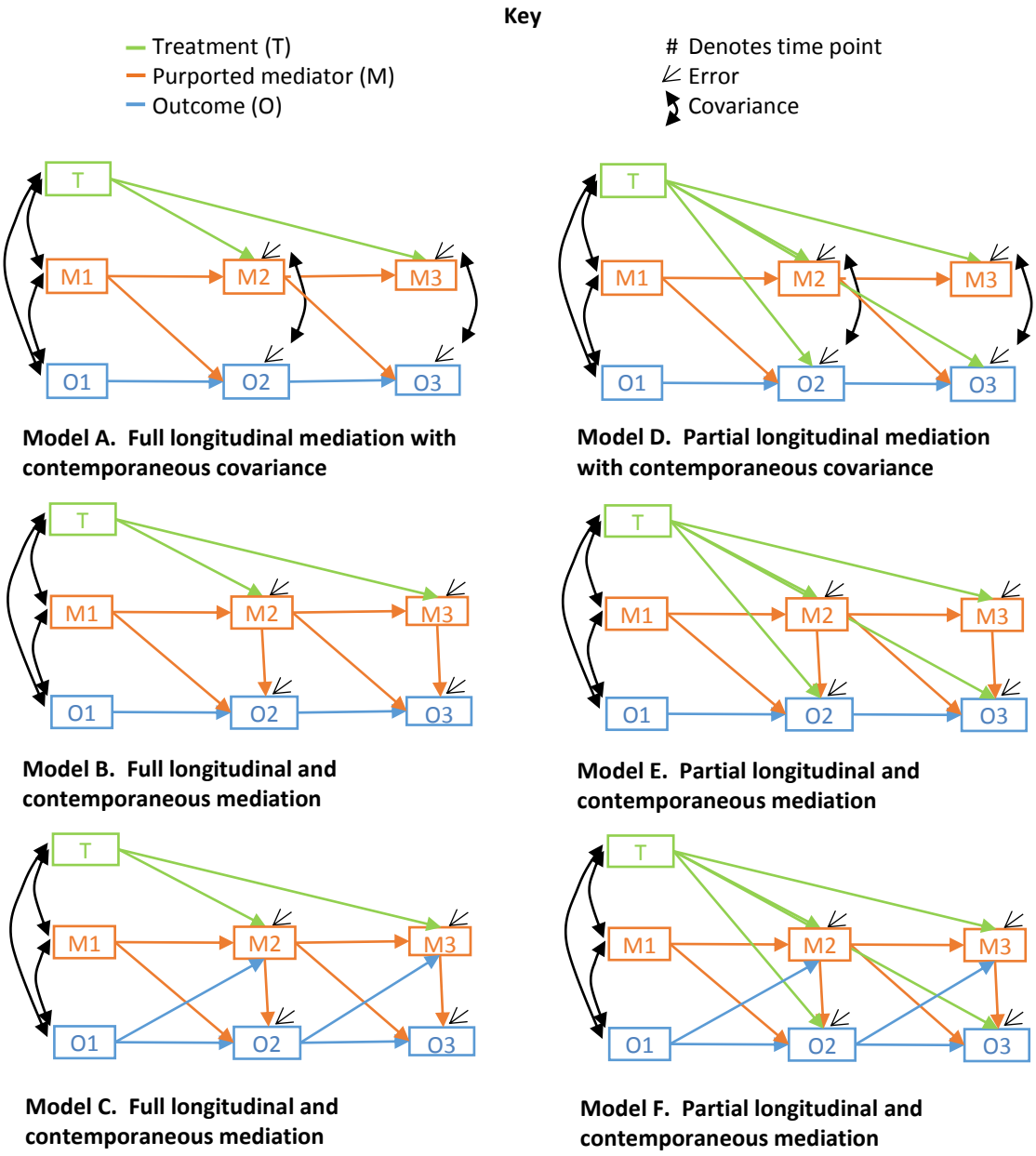


Figure 19: Different specifications of three panel autoregressive models for full and partial mediation when treatment allocation has been randomised at baseline (Models A-C, and D-F, respectively). Models A and C illustrate longitudinal mediation controlling for covariance between the purported mediator and the outcome within the same time point. Models B and D depict longitudinal and contemporaneous relationships between the purported mediator and the outcome. Models C and E demonstrate cross-lagged relationships between both the purported mediator and the outcome.

The mediated effect can be calculated using the product of coefficients method, by multiplying the coefficients on the relevant paths. There are two potential mediated effects that can be calculated in the autoregressive model. These are the "time-specific mediated effect" and the "total mediated effect" (Cole & Maxwell, 2003). The time-specific indirect effect calculates the longitudinal mediated effect at a specific time point e.g. for time point O3 this would use the paths T→M2→O3 in the models in Figure 19. In contrast, the total indirect effect calculates the mediated effect across all of the time points in the model. For Model A, both the time-specific and total indirect effects are the same as there are no other indirect paths from treatment to the outcome. However, in more complex models, such as Model E, the total effect is made up of several paths, which need to be summed (e.g. T→M2→O3, T→M2→O2, T→M3→O3, T→M2→M3→O3, T→M2→O2→O3). In partial mediation, it is also possible to calculate the indirect effect of treatment on the outcome occur via its impact on treatment at earlier time points e.g. T→O2→O3. The total effect of treatment can be calculated in a similar way, by summing across all paths across the autoregressive model.

4.1.2.2. Latent Growth Curve Models

One limitation of the autoregressive model is that by focusing on the relationship between variables at discreet time points (the covariance structure), it does not explicitly examine change over time. This means that while it can identify fluctuations in the purported mediator and the outcome, and their relationship to treatment, it cannot quantify the extent of change. Nor can it identify to what extent change in the purported mediator was related to change in the outcome, which is arguably of more

interest in the context of understanding how treatment works. Furthermore, autoregressive path coefficients incorporate information about both individual differences and within-participant change, meaning that they can be interpreted in a number of ways (Hertzog & Nesselroade, 1987). For example, high autoregressive path coefficients may indicate a lack of change in the variable, or that the variable changes in a similar way across most participants, or that prior levels determine later levels of the variables suggesting that change over time might have different rates of growth (Hertzog & Nesselroade, 1987; Selig & Preacher, 2009). Due to its focus on individual time points, autoregressive models are also insensitive to non-linear trends over time, which are reduced to greater or lesser fluctuations at different time points. The latent growth curve is an alternative SEM approach, which uses latent variables to model the initial level and subsequent change in a variable over time (Little, 2013; Muthen & Curran, 1997; Willett & Sayer, 1994). These models are essentially multilevel models, similar to the MLM described previously, but using SEM techniques (Biesanz, Deeb-Sossa, Papadakis, Bollen, & Curran, 2004; Little, 2013). As opposed to the autoregressive model, latent growth curve models examine the mean structure of the data (means and covariances).

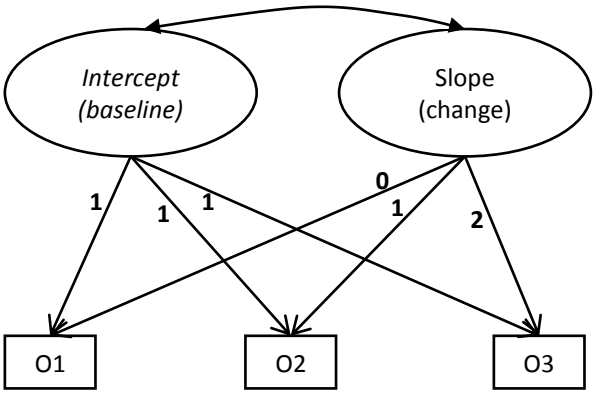


Figure 20: Example latent growth curve model for outcome O over three time points (error terms omitted)

In latent growth curve modelling the different measurements of the observed variable at different time points are modelled as indicators of two latent variables, the 'intercept' and the 'slope' (see Figure 20). The variables' paths are constrained / weighted in different ways. The intercept variable is constrained so that it represents the mean level of the variable, typically at baseline, the slope variable so that it represents the mean growth or change. For the intercept, all the paths to the different measurements at different time points are constrained to be 1. For the slope variable, the paths to the different measurements at different time points are weighted so that they represent the expected rate of change. For example, in Figure 20, the slope variable paths are constrained to be 0, 1, 2. This indicates that change is expected to increase in a linear fashion. Other constraints could be 0, 1, 1, which would indicate a plateau in the rate of change after time point 2 (i.e. a non-linear rate of growth). Alternatively, path weights can be estimated empirically. This requires one path to be constrained to be 0 (usually representing baseline) and one to be constrained to 1 (usually a point of particular interest e.g. the point at which an intervention gets implemented, or the last time point). The remaining unconstrained paths are then estimated by the programme. This can be helpful to identify the proportion of change that has occurred at each time point, when the rate of change is not known. Another way to model linear and non-linear rate of change in the model (similar to how this would be modelled in MLM) would be to model two latent slopes, a linear slope and a quadratic slope weighted as 0, 1, 2 and 0, 1, 4 respectively (Biesanz et al., 2004). This is an example of a piecewise model, modelling two distinct rates of change.

The slope reflects the mean change that has occurred between the measurements whose paths have been constrained to 0 and 1 respectively. In the linear model shown in Figure 20, the slope variable will therefore reflect the rate of change between time point 1 and 2, and also 2 and 3, because the paths are weighted similarly. In contrast, if the slope paths were constrained to 0, 0.5, 1 (also a linear model), the slope would reflect the mean change occurring between measurements 1 and 3. The intercept and slope are set so that they covary, implying the weakest assumption that baseline level of the variable may be related to rate of change. Once the initial latent growth curve has been specified, covariates can be added into the model to predict both the intercept and the slope, or in the case of time-varying covariates, the observed variables acting as indicators of the latent variables.

4.1.2.2.1. Latent Growth Curve Models for Statistical Mediation (Parallel Process Model)

Statistical mediation using latent growth curve models requires several steps (Cheong et al., 2003). First, separate latent growth curve models need to be specified for treatment, the purported mediator and the outcome. If treatment represents a one-off event, such as treatment randomisation at the start of an RCT, then treatment can be modelled as a covariate predicting the slopes of the purported mediator and the outcome, with treatment and intercepts covaried to take into account any association between treatment and the purported mediator and outcome at baseline (if present). Once the latent growth curve models for the purported mediator and the outcome have been finalised, then these models will need to be combined into one model, known as a parallel process growth model (see Figure 21).

When combining the models, the intercepts of the purported mediator and the outcome are specified to covary, representing that they may have been related at baseline. As in the univariate latent growth curve models, the intercept and slope for each variable is also specified to covary. Finally, cross-lagged paths are specified between the intercepts and slopes of the purported mediator and the outcome respectively, and a direct path from the slope of the purported mediator to the outcomes is specified. Mediated effects can be calculated using the product of coefficients method. It has been suggested that two different mediated effects can be calculated, one via the intercept of the purported mediator to the slope of the outcome ('intercept-only mediation') and one via the slope of the purported mediator to the slope of the outcome ('slope-only mediation' (von Soest & Hagtvet, 2011). Intercept-only mediation may be relevant when it is thought that baseline levels of the purported mediator may mediate later change in the outcome. This may happen when the independent variable has not been randomised and the independent variable may influence the baseline level of the purported mediator. In contrast, slope-only mediation reflects the scenario where change in the purported mediator mediates later change in the outcome. It is the latter that is of most interest in the context of understanding how treatments work.

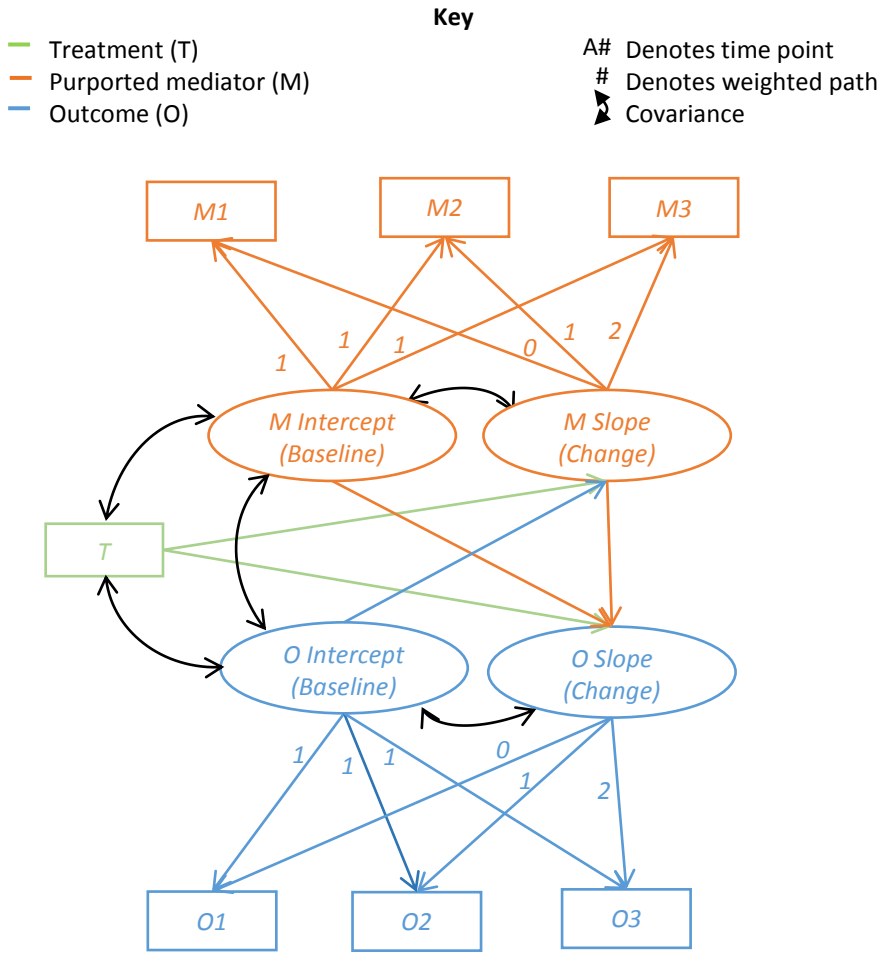


Figure 21: Parallel Process Latent Growth Curve Models for three time points (error terms omitted)

4.1.3. Statistical Power to Detect a Mediated Effects

The causal steps approach typically has low power to detect statistical mediation when this is based solely on the significance of the regression equations, meaning that they are likely to miss a genuine effect (MacKinnon et al., 2002). Whilst statistical mediation based on the calculation of a mediated effect (e.g. the product of coefficients method) is better powered, statistical tests of the mediated effect (such as the Sobel test) are also underpowered (MacKinnon et al., 2002). This is because the mediated effect is asymptotic (i.e. not normally distributed). Tests that address the asymptotic nature of the mediated effect, such as those that construct confidence intervals around the mediated effect, demonstrate better power than more traditional tests of the mediated effect (MacKinnon et al., 2002). These include resampling methods (e.g. bootstrapping), which construct confidence intervals based on multiple resamples of the data to estimate the distribution of the sample. Simulation studies estimating a mediated effect using causal steps approach suggest that bootstrapped bias-corrected confidence intervals (which adjust for the asymptotic nature of the data) are the most powerful of these methods (Fritz & Mackinnon, 2007). These studies suggest that a sample size of 148 is sufficient when a' and b' have an effect size of 0.26. When at least one path has a moderate effect size (0.39) a sample size of 115-118 is sufficient.

There is a lack of research into power to detect a mediated effect when using MLM. In structural equation modelling, it has been suggested that small sample sizes are likely to be underpowered (Wu & Jia, 2013). However, results from a simulation study suggest that a mediated effect can be detected in sample sizes as small as 125 when a'

and b' have an effect size of 0.39 (Wu & Jia, 2013). This holds even in non-normal data when up to 20% of data are missing if these issues are addressed using bootstrapping and full information maximum likelihood estimation. There has been little research into power calculations for longitudinal mediation analyses (Cole & Maxwell, 2003). Power may be particularly problematic in cross-lagged models given the number of paths that are typically need to be estimated. As complexity of the model increases, power reduces (Kline, 2011).

4.2. Modern Approaches to Statistical Mediation Analysis

Statistical mediation implies several inferences regarding causality. It implies that treatment *causes* change in the purported mediator, and that the purported mediator *causes* change in the outcome. It has been argued that the traditional approach to statistical mediation does not adequately assess these causal inferences. In traditional analyses, treatment effects are calculated as the observed difference between patients who received the intervention and those who did not (Rubin, 1974). This is arguably a proxy for what is of real interest; namely, what the patients' outcome would have been had they received treatment, compared to what it would have been had they not (Rubin, 1974). The latter represents a true quantification of the causal effect of treatment; but in practice, represents an impossible situation. Only one of these scenarios can be observed for each patient; the other scenario is unobserved. The latter is an example of a counterfactual or potential outcome. A new approach to statistical mediation is causal mediation analysis, which uses calculus to formally define and estimate causal effects. As opposed to relying on the observed associations between variables, causal mediation analysis utilises logic-based arguments to

estimate causal effects to derive both the observed and unobserved elements in the model. Whilst this technique is becoming more recognised, software has only recently become available for it and it has therefore not been widely used. Although the approach has been adapted to incorporate longitudinal data, the method is complex and computationally demanding (Daniel, De Stavola, & Cousens, 2011). It is also potentially more sensitive than traditional methods to violations of assumptions e.g. model misspecification, measurement error, unmeasured confounding (Taubman, Robins, Mittleman, & Hernán, 2009).

4.3. Analytic Strategy

As described above, there are a variety of analytic methods for undertaking statistical mediation. However, each type of analysis has its own strengths and limitations. Whilst longitudinal data represents the strongest data for investigating mediational relationships, it requires the use of more complex statistical analyses. In the context of treatment, analyses may be further complicated when treatments are multi-component, and are personalised. In RCTs that compare active treatments, interpretation may also be more difficult, particularly if change occurs simultaneously in both treatments. This is because when treatment acts as the independent variable this essentially assesses whether there is a difference between treatments in terms of the purported mediator and the outcome. This means that if there is a non-significant relationship between treatment and the purported mediator and / or the outcome the result is not easily interpretable (Doss & Texas, 2006; Hollon, DeRubeis, & Evans, 1987). Whilst it could mean that the purported mediator does not mediate the effect of treatment on the outcome, it could also reflect the fact that the treatments have

the same purported mediator, or that change in the purported mediator in one treatment is the result of another mediator (i.e. change in the purported mediator is an outcome). There is no clear guidance in the literature regarding how to undertake statistical mediation analyses under these conditions.

An analytic strategy was developed to address the issues related to the dataset described in chapter 3 that could impact on statistical mediation analysis and its subsequent interpretation. As discussed in chapter 3, this dataset consisted of a series of longitudinal mediation studies embedded within an RCT comparing enhanced cognitive behaviour therapy (CBT-E) and interpersonal psychotherapy for eating disorders (IPT-ED). Specific issues potentially affecting the analysis included;

- The comparison of two active treatments, potentially confounding treatment specific effects on the purported mediator and the outcomes.
- The use of longitudinal data using the same measure repeated at each measurement occasion necessitating more complex statistical modelling
- A lack of knowledge regarding the precise timing of change in the purported mediator and change in the outcome, meaning that the timing of measurements might not be optimal for capturing temporal precedence
- The lack of randomisation of the purported mediator, potentially leaving the relationship between the purported mediator and the outcome vulnerable to unmeasured confounding (as is usual in most mediation analyses)
- Treatment personalisation in stage three (week 5 onwards) of CBT-E, impacting on when different procedures might be implemented and when change in the purported mediator and the outcome might be observed

The analytic strategy that was developed involved a series of statistical substudies to give a multi-faceted perspective on the observed relationships between treatment, the purported mediator and the outcome. MLM and SEM were used in combination, with MLM used as a preliminary step before proceeding onto more complex models using SEM. Progress from one study to the next was dependent on whether there was evidence of statistical mediation warranting further investigation using more complex analyses.

Purported mediators of CBT-E and IPT-ED were investigated in isolation to enable accurate estimation of direct and indirect effects of treatment on the outcome. Statistical considerations particular to individual substudies (e.g. treatment personalisation) were addressed in each study chapter. Non-specific predictors and moderators of the purported mediator and the outcome were included in each study to increase precision of estimates. Potential unmeasured confounders were considered at each stage in the analysis. All analyses were undertaken in Stata 13 (StataCorp, 2013).

4.3.1. Data Preparation

4.3.1.1. Data Quality Spot Check

To assess data quality, a spot check of the data was carried out in a random sample of 26 patients (20% of the overall sample); it examined data entry for the Eating Questionnaire (EQ), Regular Eating Rating Scale (RERS), Binge Eating Rating Scale (BERS), and the Interpersonal Interview (IPI) (measures that were used in the studies in this thesis; see chapter 3 for more details). The Eating Disorder Examination (EDE), Beck Depression Inventory II (BDI-II) and Rosenberg Self-Esteem Scale (RSES) were

excluded from the spot check as data cleaning for these measures was done separately as part of the main trial analyses. All intervention record sheets (see chapter 3) were checked at the point of data entry.

Database entries were compared to hard copy records to assess accuracy of data entry, and to identify any systematic errors in data entry. An error rate was calculated for each patient based on the number of items entered incorrectly for that patient divided by the total number of questionnaire items (see Table 10 for more detail), and then an average error rate calculated across the spot check sample.

Table 10: Breakdown of number of items in dataset by measure and number of repeated measurements

Measure	Number of Items per Measure	Number of Times Repeated	Total Number of Data Entry Items
IPI	16^{*} / 10^{+}	4	46
RERS	1	5	5
BERS	1	5	5
EQ	31	17	527
All			583

Notes. IPI=Interpersonal Interview, * at initial interview, $^{+}$ at subsequent interviews, RERS=Regular Eating Rating Scale, BERS= Binge Eating Rating Scale, EQ=Eating Questionnaire.

The mean number of errors per case was 2.19 items (SD: 3.86), representing an average error rate of 0.38% per patient, with over a quarter of patients (7 patients, 27% of the spot check sample) having no errors on these measures, and two thirds (18 patients, 69% of the spot check sample) having an error rate of less than 0.20%. No systematic errors were identified.

4.3.1.2. Routine Data Cleaning

In addition to the initial spot check, all data were further assessed at the start of each study to assess and correct for out of range values and problems with missing data, such as systematic missing values. One major issue was discovered. An unusually high

amount of missing data were found on the RERS and BERS at week 3 of treatment in IPT-ED compared to CBT-E, and later weeks of treatment. This suggested that these data were not missing at random. Further investigation indicated that daily self-monitoring was terminated early in IPT-ED in week 5 (not week 3) for a number of patients. The data were corrected so that the data that had been entered incorrectly at week 5 were attributed to the correct time points (weeks 3 and 4). As the week 5 data were not missing at random, it was decided that week 5 data from both treatments would be excluded from all analyses involving these measures. No other problems with missing data were detected, with the majority of missing data being due to attrition. Data were assumed to be missing at random, once treatment week was taken into account.

4.3.1.3. Assumption Checking

Distributional assumptions were checked by visually inspecting histograms and q-q plots of the purported mediators and outcomes (further details are provided in each study chapter). Bootstrapping based on 1000 resamples was used to construct bias-corrected 95% confidence intervals to correct for non-normality in the data for the final models.

4.3.2. Study A: A Preliminary Investigation using Multilevel Modelling for Statistical Mediation

The causal steps approach used in MLM provides a good framework for understanding the basic relationships between treatment, the purported mediator and the outcome (Dunn et al., 2015). MLM has greater model stability and smaller sample size

requirements than SEM and therefore was a natural preliminary step prior to modelling the direct relationships between treatment, the purported mediator and the outcome using SEM.

MLM was undertaken in four steps, based on the updated causal steps approach (Kraemer et al., 2002; MacKinnon et al., 2000, 2002; Shrout & Bolger, 2002);

- Step 1: Assessing the impact of treatment on the purported mediator
- Step 2: Assessing the impact of treatment on the outcome
- Step 3: Assessing the relationship between the purported mediator and the outcome
- Step 4: Assessing the impact of treatment on the outcome when accounting for the purported mediator

The MLM undertaken in these analyses, examined trends over time during treatment. Baseline differences were not constrained to be zero due to lack of randomisation of the purported mediator. Treatment week was entered as non-specific predictor, with baseline coded as 0, and subsequent weeks as 1, 2, 3 etc. A quadratic term was included for treatment week to assess for non-linear change in the purported mediator and the outcome. An interaction was included between treatment and treatment week to explore the effect of treatment over time. Treatment week was tested as a random effect in all analyses.

Potential (a priori determined) non-specific predictors and moderators of the purported mediator and the outcome were controlled for in the analysis to increase estimation precision. Non-specific predictors were added as main effects only. Moderators were added as main effects and as an interaction with treatment. These

variables were selected based on research undertaken by Cooper and colleagues, which examined non-specific predictors and moderators of treatment outcome at 12 months in the main trial (Z. Cooper et al., 2016) and in the shorter term (E. Allen, personal communication, 25 September 2015). They included the following non-specific predictors assessed at baseline;

- importance of shape
- duration of eating disorder
- age
- eating disorder diagnosis (bulimia nervosa vs. eating disorder not otherwise specified)
- weight concern
- depression
- self-esteem

The following moderator was also assessed;

- self-esteem status (high self-esteem vs. low self-esteem)

Where relevant, additional study specific non-specific predictors and moderators were also assessed in each chapter, based on theory

All binary non-specific predictors / moderators were dummy coded as 0 and 1, and continuous non-specific predictors / moderators were centred around the sample mean at baseline (rounded to the nearest integer in the case of variables denoting years) so that the model constant represented baseline status (in those groups coded as 0). Non-specific predictors / moderators were retained in the model based on a

combination of forwards and backwards elimination. Each variable was tested initially in isolation by adding it into the basic model, firstly on its own, and then as trends over time by adding an interaction with treatment week (both its linear and quadratic trends). Quadratic relationships between the continuous non-specific predictors and the outcome were also tested by squaring the original variable using an interaction with itself. Variables (and interactions) that significantly predicted the outcome (based on their bias-corrected confidence intervals) were then included in a final model.

In steps 3 and 4, the purported mediator was centred at the patient's baseline score, so that the constant still reflected the mean at baseline in those groups coded as 0. The relationship between the purported mediator and the outcome was tested both contemporaneously (i.e. as their relationship within the same treatment week), and as a lagged relationship (i.e. as the relationship between the purported mediator at the previous week and the outcome the following week). This was done to assess the temporal relationships between these variables. The relationship between the purported mediator and the outcome over time was also examined using smoothed-mean line graphs (which predicts the relationship between the two variables using least-squares estimation).

For the final models presented in this thesis, all non-significant predictors were removed, with the exception of treatment, the interaction between treatment allocation / intervention status and treatment week (linear trend only), and the purported mediator, as the primary relationships of interest. Models were trimmed using backwards elimination. As MLM was used only as a preliminary step in model building, mediated effects were not calculated. Models were estimated using

restricted maximum likelihood estimation. These final models were used to inform modelling decisions in studies B, C and D.

4.3.3. Study B: An In-Depth Analysis using Autoregressive Structural Equation Modelling for Statistical Mediation

Autoregressive SEM was used to assess the relationships between treatment, the purported mediator and the outcome at each treatment week. Autoregressive paths were specified for the purported mediator and the outcome, with error terms for the same repeated measure were specified to co-vary to take into account shared measurement error. Stability of the autoregressive path coefficients was examined to assess whether the autoregressive path remained stable during treatment. This was done using Wald tests following model estimation. Stationarity was not assumed as it was expected that variables would change during the course of treatment.

A partial mediation model was specified, with treatment predicting both the purported mediator and the outcome from the first week following initiation of treatment procedure of interest (i.e. from when the treatment could have begun affecting the purported mediator / outcome). Longitudinal and contemporaneous paths were specified between the purported mediator and the outcome at each time point to model relationships over time and within the same treatment week (Model E in Figure 9 of Chapter 3). This model was chosen as it allowed modelling statistical mediation under conditions of non-optimal timing of measurement (where change in the purported mediator and the outcome had both occurred before the first measurement occasion).

Non-specific predictors / moderators identified in Study A as having significant effect on the purported mediator / outcome were included as exogenous variables, covaried at baseline with treatment, the purported mediator and the outcome. Paths were then specified from each non-specific predictor / moderator to the relevant variable at each treatment week (irrespective of whether a significant linear / quadratic trend had been found in study A) in order to stabilise the model.

Model estimation used full information maximum likelihood to take into account missing data. An unstructured covariance matrix was specified so that covariance and variances were estimated freely, without any assumptions imposed on these variables and their relationships. Both unstandardised estimates (in the variable metric) and standardised estimates (as an estimate of effect size) are reported. Mediated effects were calculated using the 'teffects' command in Stata, which calculated the indirect effect of treatment on the outcome at each time point using the product of coefficients method. This included effects achieved both via the purported mediator and those achieved via the autoregressive path (i.e. through earlier effects on the outcome). Whilst this means that this was not a 'pure' assessment of the mediated effect (i.e. the effect achieved *only* via the purported mediator), it reflected a more accurate picture of the effect of treatment on the outcome over time. Bias-corrected bootstrapped confidence intervals were constructed around the mediated effect using 1000 resamples.

4.3.4. Study C: Testing an Alternative Model using Autoregressive Structural Equation Modelling for Statistical Mediation

Study C was used to assess the specificity of the relationships between treatment, the purported mediator and the outcome, by testing an alternative model. It followed the same procedure as study B, but, contrary to the hypothesised model, the outcome was modelled as an alternative mediator of the effect of treatment on the purported mediator (as an alternative outcome). Study C was only undertaken if studies A and B supported the purported mediator.

4.3.5. Study D: An Investigation into the Relationship between Change in the Mediator and Change in Outcome using Latent Growth Curve Modelling

A parallel process latent growth curve model was used to assess the relationship between change in the purported mediator and change in the outcome during the period studied. Separate growth curves were specified for the purported mediator and the outcome and these were then combined into one model. Both linear and non-linear trends were assessed in each variable, by constraining the slope path of the first week to be 0 (indicating baseline) and the slope path of the last week of the period studied to be 1. The intervening weeks were then left unconstrained so that they were freely estimated by the programme in order to assess the shape of the growth curve. This approach was selected as it is a less demanding than modelling a separate quadratic slope. This pattern of weighting was chosen so that the slope reflected the mean amount of change occurring in each variable during the period studied, thereby

preventing potential difficulties with model interpretation should there be different rates of change in the purported mediator and the outcome.

A partial mediation model was specified, with treatment predicting both the intercept and slope of the purported mediator and the outcome. Paths were specified from the intercept and slope of the purported mediator to the slope of the outcome. Unlike the original parallel process latent growth curve model of Cheong and colleagues (Cheong et al., 2003), no path was specified between the intercept of the outcome and the slope of the purported mediator. This was done so that the model replicated previous model building choices regarding the exclusion of feedback loops in studies B and C. The mediated effect was calculated manually using the product of coefficients method by multiplying the coefficient of the path from treatment to the slope of the purported mediator and the coefficient of path from the slope of the purported mediator to the slope of the outcome. Bias-corrected bootstrapped confidence intervals were then constructed around the mediated effect using 1000 resamples.

Study D was only undertaken if studies A, B and C supported the purported mediator.

4.3.6. Criteria for Accepting a Purported Mediator as a Potential Mechanism of Action of Treatment

A purported mediator was accepted as a potential mechanism of action of treatment if it met the following criteria;

- In study A;

    - change in the purported mediator occurred before change in the outcome (i.e. temporal precedence of the purported mediator could be established)
    - inclusion of the purported mediator diminished the effect of treatment on the outcome
- In study B:
    - the model was an acceptable fit to the data (chi-square > .05, CFI > .95 and RMSEA <.08)
    - the strength and direction of the model path coefficients were consistent with mediation via the purported mediator
    - there was a significant mediated effect via the purported mediator based on the bootstrapped bias-corrected confidence intervals
- In study C:
    - The alternative model was *not* an acceptable fit to the data (chi-square <.05, CFI < .95 and RMSEA >.08)
    - the strength and direction of the alternative model path coefficients were *inconsistent* with mediation via the alternative mediator
    - There was *not* a significant mediated effect via the alternative mediator based on the bootstrapped bias-corrected confidence intervals
- In study D:
    - the model was an acceptable fit to the data (chi-square > .05, CFI > .95 and RMSEA <.08)

    - the strength and direction of the model path coefficients were consistent with mediation via the purported mediator
    - there was a significant mediated effect via the purported mediator based on the bootstrapped bias-corrected confidence intervals
- Across all studies, results were consistent regarding the purported mediator mediating the effect of treatment on the outcome.

Where change in the purported mediator could not be disentangled from change in the outcome (i.e. where measurement timing has not been optimal), but all other criteria was met, results were viewed as inconclusive.

Chapter 5. Study 1: Regularity of Eating as a Mediator of the Effect of Enhanced Cognitive Behaviour Therapy on Frequency of Binge Eating

5.1. Background

As discussed in chapters 1 and 2, many eating disorders are characterised by both under eating and over eating, although the relative balance of these behaviours vary across individuals and eating disorder diagnoses. There is evidence to suggest that the way that people with eating disorders attempt to control their eating (e.g. reducing caloric intake by reducing the amount of food eaten, avoiding certain types of food that they perceive to be 'bad' (e.g. high fat, high sugar, carbohydrates) or skipping meals) leaves them potentially vulnerable to binge eating.

In enhanced cognitive behaviour therapy (CBT-E) a core intervention in stage 1 of treatment is the "regular eating" procedure, which is usually implemented in the second week of treatment in session 2 or 3 (see Chapter 3 for treatment timings). Patients are informed about the value of establishing a regular eating pattern of three meals plus two or three snacks a day with ideally no more than four hours gap between. Outside of these times, patients are recommended to refrain from eating, and strategies to resist the urge to binge are discussed (Fairburn, Cooper, Shafran, et al., 2008b). Patients are advised that there is no need to change what they eat at this

stage in treatment. It is recommended that all the meals and snacks are planned in advance and that their eating should not be followed by compensatory behaviour.

5.2. Aim/Hypothesis

The cognitive behavioural model of eating disorders proposes that rigid dietary restraint leaves the patient vulnerable to binge eat due to the strict, inflexible manner in which patients typically implement dieting. It therefore hypothesises that by increasing adherence to a regular pattern of eating this will lead to a reduction in the patient's frequency of binge eating (see Figure 22).

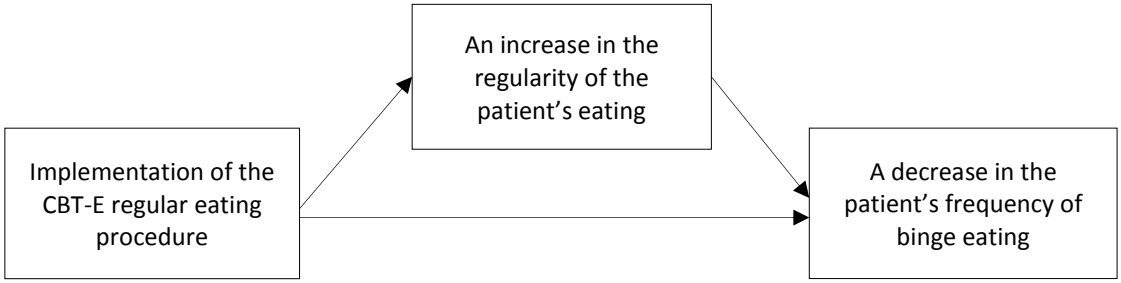


Figure 22: Illustration of the purported mediational relationship between enhanced cognitive behaviour therapy (CBT-E), regularity of eating and frequency of binge eating

In contrast, interpersonal psychotherapy for eating disorders (IPT-ED) does not directly address or refer to the patient's eating pattern. Whilst frequency of binge eating might decrease in IPT-ED via improvements in interpersonal functioning (thus reducing the number of adverse events triggering binge eating), such change would likely occur more slowly than in CBT-E as interpersonal problems do not start to be addressed until a few weeks into therapy.

The aim of this study was to assess whether change in patients' regularity of eating mediates change in their frequency of binge eating to a greater extent in CBT-E compared to IPT-ED. It was hypothesised that patients in CBT-E would increase their

regularity of eating more rapidly than in IPT-ED and that this increase would mediate change in frequency of binge eating in CBT-E but not IPT-ED. It was also hypothesised that the effect of this intervention might differ between patients who had experienced binge eating in the month prior to starting treatment and those who had not, therefore this was tested as a moderator (see analytic considerations for more details).

5.3. Method

5.3.1. Experimental Procedure

Patients in each arm of the trial were asked to record their eating habits and related thoughts, feelings and events using a monitoring record (see Appendix 12 and Appendix 13) which were collected at each therapy session. This occurred for the first 4-5 weeks of IPT-ED as part of the research study and on an ongoing basis in CBT-E as self-monitoring is routinely part of treatment. However, as the majority of changes attributable to the regular eating procedure are thought to occur in the first few weeks after the introduction of the procedure (Fairburn, Cooper, Shafran, et al., 2008b) (weeks 1-4 inclusive), only the first 4 weeks of these data were used for the purpose of this study. These records were examined to assess the regularity of the patient's eating and the presence of any binge eating.

These records were subsequently photocopied and collated by intervention week (as opposed to chronological week[4]), with date and treatment identifying information removed, so that they could be assessed by raters blind to intervention week and

[4] Due to the nature of real-world therapy, the spacing of therapy sessions varied based on availability of patients and therapists. Therefore appointments did not necessarily occur within the space of a chronological week. It was therefore decided to base collation of self-monitoring records on the "dose" of therapy received i.e. number of sessions received.

treatment allocation. The records were then assessed by two independent raters to determine the patient's adherence to a regular eating pattern and their frequency of binge eating using the 'Regular Eating Rating Scale (RERS)' and 'Binge Eating Rating Scale (BERS)' (see Chapter 3). The records were first rated separately by the two raters, and then they consulted to agree upon a final score.

5.3.2. Study Measures

5.3.2.1. Regularity of Eating

Regularity of eating was assessed using the RERS, which was developed for the present study (see Appendix 3 and Appendix 4). It assessed the patient's adherence to a regular eating pattern (based on their monitoring records) of 3 meals plus 2 or 3 snacks (with no greater than 4 hour gaps between episodes of eating) during the past week. Regularity of eating was rated on a scale of 0-6, where 0 indicated 'no regular eating' and 6 indicated 'marked adherence'.

5.3.2.2. Frequency of Binge Eating

Frequency of binge eating was assessed using the BERS, which was developed for the present study (see Appendix 4 and Appendix 5). It assessed average number of binge eating episodes per day over a week. These episodes were defined as eating a substantially larger amount of food than most people would eat during a 2 hour period. It was not possible to assess the presence of a sense of loss of control at the time. For the purposes of this study this number was converted into a weekly average (by multiplying the daily average by 7) to have a more clinically meaningful outcome.

5.3.3. Analytic considerations

5.3.3.1. Potential Confounders

Potential unmeasured confounders that were considered were; depression, distressing interpersonal events occurring outside of treatment and the violation of the patients' dietary rules. All three were potentially purported mediators of the effect of treatment on the outcome, and therefore would have been excluded from the analysis. Although depression was not measured, the influence of clinical depression on these variables were controlled for as those patients requiring treatment for depression were treated before entering the trial, and anti-depressant medication was maintained during treatment. Those who became depressed during treatment were put on hold while the depression was treated.

5.3.3.2. Presence of Binge Eating as a Potential Moderator

Presence of binge eating in the month prior to starting treatment was tested as an additional moderator in this study and was based on the frequency of binge eating reported in the Eating Disorder Examination undertaken prior to commencing treatment (see Chapter 3). A binary variable was created to identify patients who had experienced an episode of binge eating in the month prior to starting treatment and patients who had not. Presence of binge eating was tested as a moderator as it was conceived possible that patients who did and did not binge eat might respond differently to the regular eating procedure. In study 1a (A Preliminary Investigation using Multilevel Modelling for Statistical Mediation), presence of binge eating was tested as a moderator using a simple interaction between treatment allocation and presence of binge eating.

5.4. Results

5.4.1. Non-Normality and Missing Data

Examination of the data suggested non-normality, particularly in frequency of binge eating, which was positively skewed throughout treatment due to a high number of patients in the sample not bingeing. In contrast, regularity of eating was negatively skewed in the CBT-E group at baseline, suggesting higher levels of regularity of eating in this treatment arm. Missing data was mostly due to attrition (although see chapter 4 for a more detailed discussion of these data).

5.4.2. Study 1a: A Preliminary Investigation using Multilevel Modelling for Statistical Mediation

5.4.2.1. Step 1: Impact of Treatment on the Purported Mediator (Regularity of Eating)

Regularity of eating (purported mediator) changed little in IPT-ED but increased in CBT-E (see Figure 23 and Figure 24). There was a significant interaction between treatment allocation and treatment week. In CBT-E regularity of eating increased significantly by 1.12 points ($p < .001$) more each week than IPT-ED (see graph 1, Figure 25). This was qualified by a significant quadratic of trend of -0.23 (p=.001), suggesting that change in regularity of eating stabilised in weeks 2/3 of treatment in CBT-E. In contrast to CBT-E, regularity of eating did not change significantly in IPT-ED during this period. Presence of binge eating in the month prior to starting treatment did not significantly moderate the effect of treatment and its interactive effect with treatment was therefore

dropped from the final model. See Appendix 14 for Stata code and Appendix 15 for full model output.

5.4.2.2. Step 2: Impact of the Treatment on the Outcome (Frequency of Binge Eating)

Frequency of binge eating (outcome) decreased in both treatments; however, the frequency of binge eating declined earlier in CBT-E compared to IPT-ED (see Figure 23 and Figure 24). There was a significant interaction between treatment allocation and treatment week. In CBT-E frequency of binge eating decreased by -1.40 (p = .007) episodes of binge eating each week (see graph 2, Figure 25). This was qualified by a significant quadratic trend of 0.43 (p = .006), suggesting that the rate of decline of frequency of binge eating slowed in CBT-E compared to IPT-ED. In contrast to CBT-E, in IPT-ED the frequency of binge eating increased initially by 0.75 (p = .032) episodes of binge eating each week. This was qualified by a significant quadratic trend of -0.25 (p = .035), suggesting that frequency of binge eating started to decline by week 4 of treatment. Presence of binge eating in the month prior to starting treatment did not significantly moderate the effect of treatment and its interactive effect with treatment was therefore dropped from the final model. See Appendix 14 for Stata code and Appendix 16 for full model output.

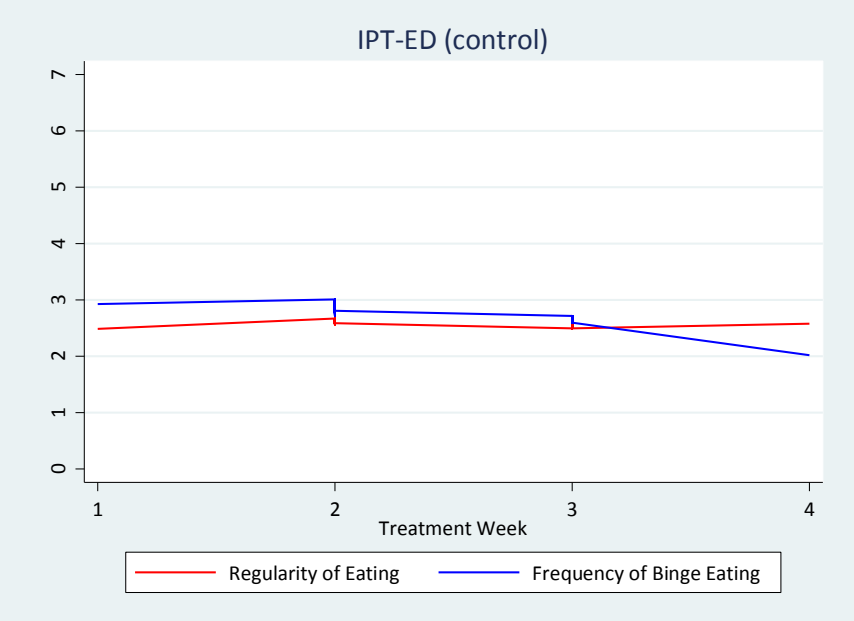


Figure 23: Line graph plotting the observed means (smoothed curve) of regularity of eating and frequency of binge eating in IPT-ED

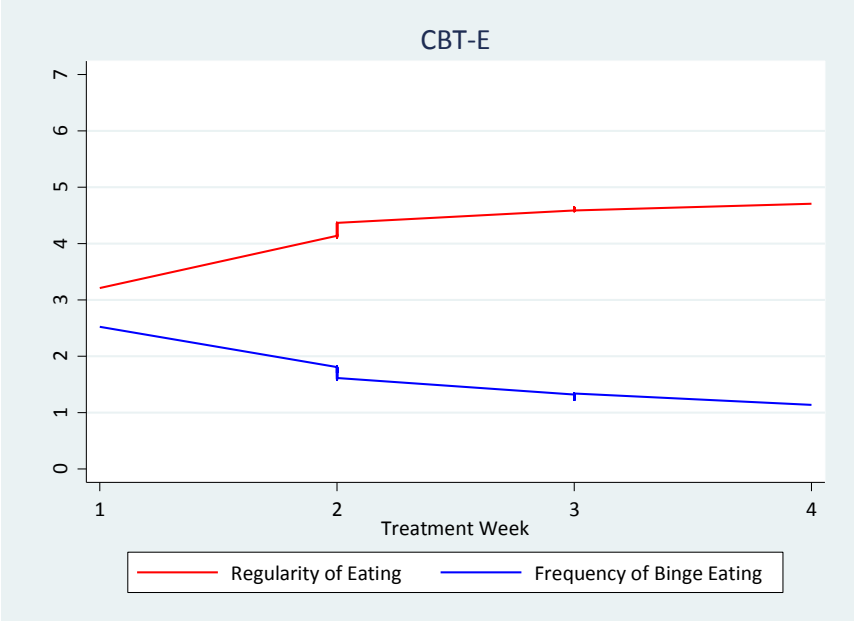


Figure 24: Line graph plotting the observed means (smoothed curve) of regularity of eating and frequency of binge eating in CBT-E

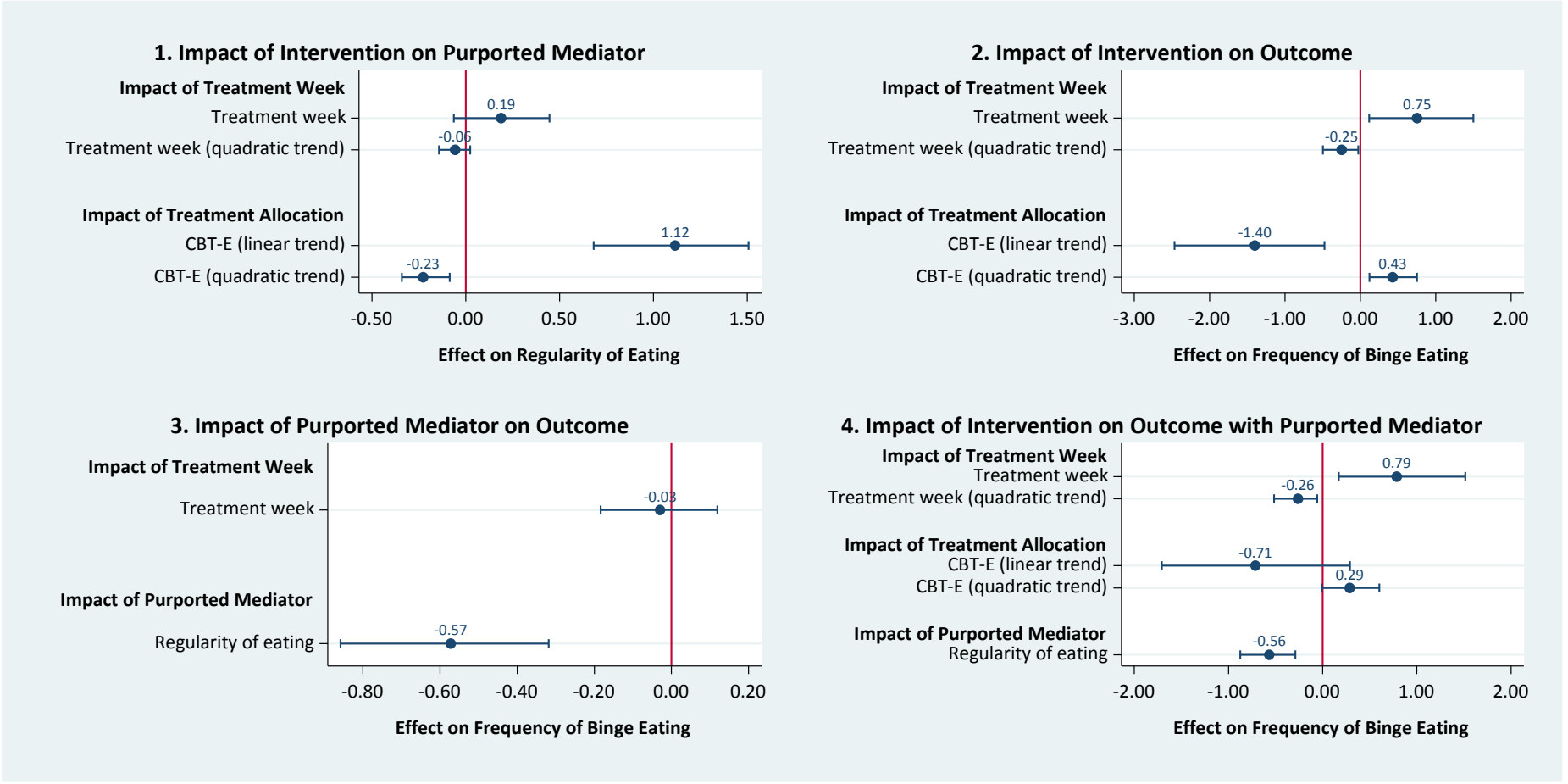


Figure 25: Coefficient plots with 95% bias-corrected confidence intervals (1000 bootstrap resamples) comparing the effect of the intervention (CBT-E) on regularity of eating and the frequency of binge eating when regularity of eating was / was not included as a mediator. Impact of treatment allocation represents an interaction between treatment allocation and treatment week. Controlling for the effects of eating disorder diagnosis, depression, weight concern, presence of binge eating in the month prior to starting treatment, and age at baseline.

5.4.2.3. Step 3: Relationship between the Purported Mediator (Regularity of Eating) and the Outcome (Frequency of Binge Eating)

Regularity of eating (purported mediator) was negatively associated with frequency of binge eating (outcome), with a one point increase in regularity of eating being associated with -0.57 ($p < .001$) fewer binge eating episodes each week (see graph 3, Figure 25). There were clear differences in the trajectories of these two variables between treatments (see Figure 23 and Figure 24). In CBT-E, change in regularity of eating and frequency of binge eating occurred simultaneously; with increases in regularity of eating mirrored by decreases in frequency of binge eating. There was no clear temporal precedence. In contrast, regularity of eating changed little in IPT-ED, and change in frequency of binge eating appeared to be unrelated to regularity of eating. See Appendix 14 for Stata code and Appendix 17 for full model output.

5.4.2.4. Step 4: Impact of Treatment on the Outcome (Frequency of Binge Eating) when Accounting for the Purported Mediator (Regularity of Eating)

When including treatment allocation and regularity of eating (purported mediator) in the same model, there was no longer a significant interaction between treatment allocation and treatment week ($p = .170$). The impact of CBT-E on the outcome (frequency of binge eating) compared to IPT-ED decreased from -1.40 episodes of binge eating to -0.71 (see graph 3, Figure 25). This represents a 51% decrease in effect. This suggests that regularity of eating explained some of the variance in outcomes between the two treatments, supporting the hypothesis that regularity of eating is a potential mediator of the effect of treatment on the frequency of binge eating in CBT-. See Appendix 14 for Stata code and Appendix 18 for full model output.

A lagged relationship between the purported mediator and the outcome (i.e. the relationship between the purported mediator at the previous week and the outcome the following week) was also tested. This was done to assess the temporal relationships between these variables. This lagged model found little evidence of a relationship between earlier measures of regularity of eating (purported mediator) and the outcome (frequency of binge eating).

5.4.3. Study 1b: An In-Depth Analysis using Autoregressive Structural Equation Modelling for Statistical Mediation

To verify the results of study 1a, namely that there was an indirect relationship between treatment and frequency of binge eating via regularity of eating in CBT-E, an autoregressive structural equation model was used to directly model the relationships between treatment, frequency of weighing and concern about weight. Presence of binge eating in the month prior to starting treatment was not included as a moderator as preliminary analyses using multilevel modelling in study 1a did not support it as a moderator.

5.4.3.1. Effect of Treatment on the Purported Mediator (Regularity of Eating)

There was a significant difference between treatments, with regularity of eating (purported mediator) increasing in CBT-E in week 2 by 1.21 points more than in IPT-ED (see Figure 26), although the strength of relationship of prior regularity of eating to regularity of eating at later time points did not vary significantly during this period; Chi2(4)=8.56, *p*=.073. This suggests that CBT-E significantly increased the purported mediator (regularity of eating) during this period.

5.4.3.2. Effect of Treatment on the Outcome (Frequency of Binge eating) when Accounting for the Purported Mediator (Regularity of Eating)

There were no significant differences between treatments in terms of frequency of binge eating (outcome) during the first three weeks of treatment when accounting for the purported mediator (regularity of eating) (see Figure 26). However, by week 4 there was a significant difference between the two treatments in frequency of binge eating, with the difference between CBT-E and IPT-E decreasing by 0.95 at week 4. Examining the model as whole this appears to reflect the fact that at week 4 the frequency of binge eating began to decrease in IPT-ED (rather than reflecting an increase in frequency of binge eating in CBT-E). However, the strength of relationship of prior frequency of binge eating to frequency of binge eating at later time points remained more or less stable during this period; Chi2(2)=2.35, *p*=.309. This suggests that treatment allocation did not have a significant effect on the outcome (frequency of binge eating) when the purported mediator (regularity of eating) was accounted for, at least in the first three weeks of treatment.

5.4.3.3. Indirect Effect of Treatment on the Outcome (Frequency of Binge eating) via the Purported Mediator (Regularity of Eating)

There was a significant relationship between the mediator (regularity of eating) and the outcome (frequency of binge eating) at all time points, with two distinct trends being observed (see Figure 26). Firstly, there was a significant negative association between the purported mediator (regularity of eating) and the outcome (frequency of binge eating) each week, with greater regularity of eating being associated with a lower frequency of binge eating within the same treatment week. Secondly, a positive

relationship was observed between the purported mediator (regularity of eating) and the outcome (frequency of binge eating) at later weeks, with higher levels of regularity of eating predicting increased binge eating the following week. Despite this trend, a negative relationship was found overall, as this positive relationship was subsumed by the negative relationship between these variables. There was evidence of an indirect effect of CBT-E decreasing frequency of binge eating via regularity of eating (see Table 11). This suggests that change in regularity of eating was related to change in frequency of binge eating in CBT-E i.e. change in the purported mediator (regularity of eating) mediated the effect of the intervention on the outcome (frequency of binge eating). See Appendix 19 for Stata code and Appendix 20 for full model output.

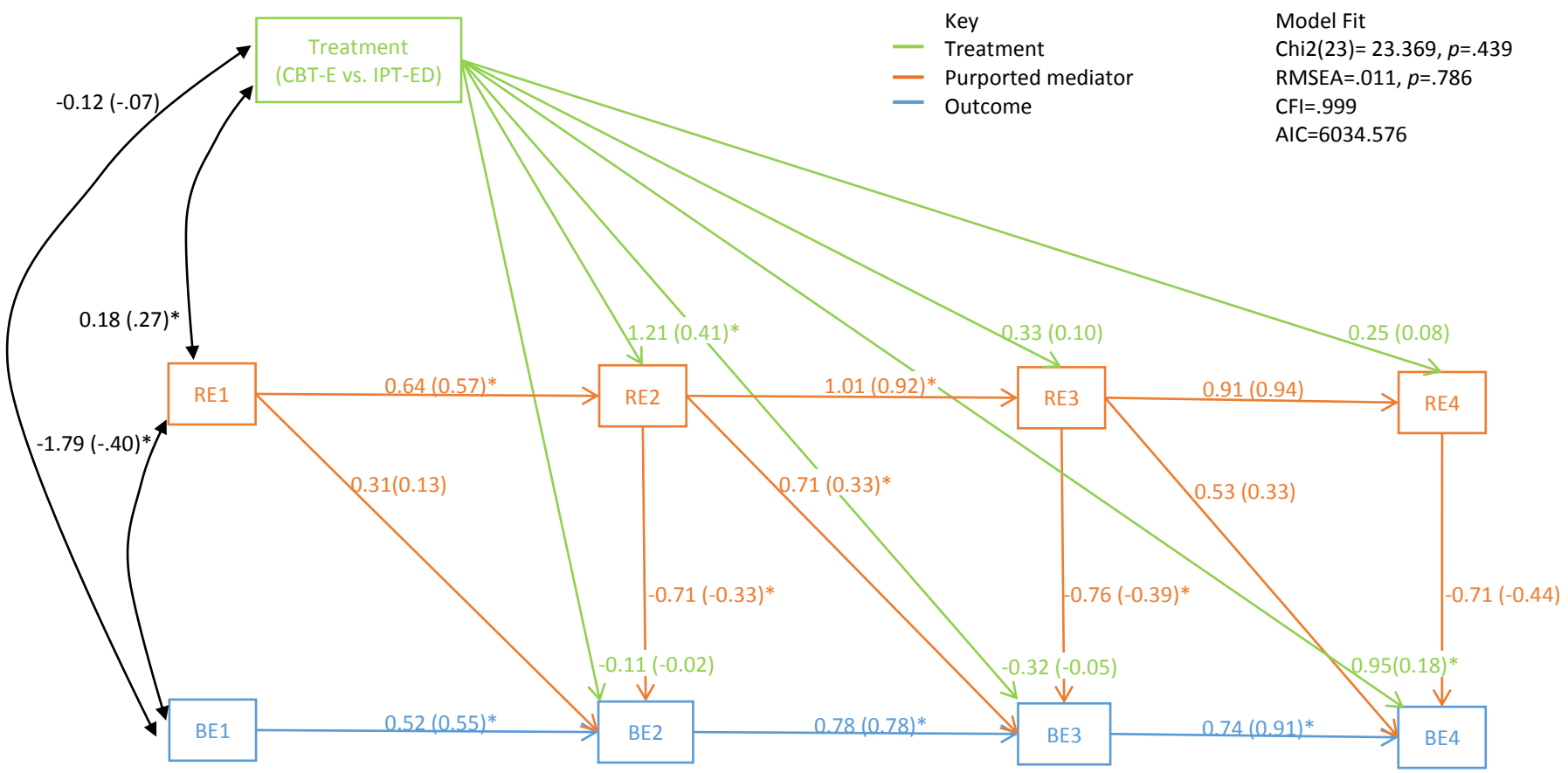


Figure 26: Autoregressive structural equation model of the impact of treatment on frequency of binge eating (BE) via regularity of eating (RE). Numeral indicates treatment week. Controlling for shared measurement variance, depression, eating disorder diagnosis, presence of binge eating in the month prior to starting treatment, age and weight concern at baseline (standardised estimates in brackets). Regularity of eating denoted as 'RE' and frequency of binge eating denoted as 'BE', where numeral indicates treatment week. Significance denoted by * based on bootstrapped bias-corrected confidence intervals (1000 resamples). Missing data accounted for using full information maximum likelihood estimation.

Table 11: Comparison of the cumulative indirect and total effects of the intervention (CBT-E) on frequency of binge eating (BE) each week controlling for shared measurement variance, depression, eating disorder diagnosis, presence of binge eating in the month prior to starting treatment, age and weight concern at baseline

Impact of Treatment Allocation on Frequency of Binge Eating	Treatment Week	B	SE	Bias-Corrected 95% CI of B Coefficient		β
				Lower	**Upper**	
Indirect effects	2	-0.86	0.37	-1.65	-0.27	-0.14
	3	-1.08	0.51	-2.10	-0.17	-0.17
	4	-1.40	0.55	-2.64	-0.48	-0.27
Total effects	2	-0.97	0.44	-1.90	-0.17	-0.15
	3	-1.39	0.53	-2.61	-0.46	-0.22
	4	-0.45	0.42	-1.34	0.31	-0.09

Notes. B=unstandardized estimate, SE=standard error, CI= confidence interval, β=standardised estimate. CIs are bias-corrected based on 1000 bootstrap resamples. Indirect and total effects of treatment incorporate transmission of treatment effects via autoregressive paths (rather than each time point in isolation).

5.4.4. Study 1c: Testing an Alternative Model: Frequency of Binge Eating (the Outcome) as Potential Mediator of Change in Regularity of Eating (the Purported Mediator)

To further test regularity of eating as a purported mediator of the effect of CBT-E on frequency of binge eating (the outcome) an alternative model was fitted to the data. This model tested an alternative hypothesis; namely that change in regularity of eating was better explained by the impact that treatment had on frequency of binge eating (the outcome) than the other way around. This was particularly relevant given that temporal precedence of the purported mediator (regularity of eating) had not been supported in studies 1a and 1b. Frequency of binge eating was tested as an 'alternative mediator' of the effect of CBT-E, and regularity of eating was treated as an alternative outcome.

5.4.4.1. Effect of Treatment on Frequency of Binge Eating (as an Alternative Mediator)

In contrast to study 1b, there was a significant difference between treatments in frequency of binge eating (alternative mediator), with patients in CBT-E showing a decrease in their frequency of binge eating in week 2 by -1.03 more than patients in IPT-ED (see Figure 27). However, the strength of relationship of prior frequency of binge eating to frequency of binge eating at later time points remained more or less stable during this period; Chi2(2)=1.59, p=.453. This suggests that treatment significantly decreased the alternative mediator (frequency of binge eating) during this period.

5.4.4.2. Effect of Treatment on the Regularity of Eating (as an Alternative Outcome) when Accounting for Frequency of Binge Eating (as an Alternative Mediator)

Similar to study 1b, there was a significant difference between treatments in regularity of eating (alternative outcome), with patients in CBT-E showing an increase in their regularity of eating in week 2 by 1.13 points more than patients in IPT-ED (see Figure 27). The strength of relationship of prior regularity of eating to regularity of eating at later time points varied significantly during this period; Chi2(2)=11.95, *p*=.003. This suggests that treatment allocation had a significant effect on regularity of eating (as an alternative outcome) even when accounting for the potential influence of frequency of binge eating as an alternative mediator.

5.4.4.3. Indirect Effect of Treatment on Regularity of Eating (as an Alternative Outcome) via Frequency of Binge Eating (as an Alternative Mediator)

Similar to study 1b, there was significant relationship between the alternative mediator (frequency of binge eating) and the alternative outcome (regularity of eating) at all time points (see Figure 27). As in study 1b both negative and positive associations between these variables were observed during this period. Whilst there was a significant indirect effect for CBT-E increasing regularity of eating (alternative outcome) via frequency of binge eating (alternative mediator) at week 2, in contrast to study 1b the effect size for this indirect effect was minimal at β=0.04 (see Table 12). However, a moderate indirect effect of treatment was observed at later weeks. This suggests that change in change in regularity (alternative outcome) was only minimally related to change in frequency of binge eating (alternative mediator) in week 2 of

treatment. The presence of an stronger indirect effect of treatment on regularity of eating (alternative outcome) at later weeks was likely related to the direct effect of treatment on regularity of eating at week 2 rather than frequency of binge eating (alternative mediator). See Appendix 21 for Stata code and Appendix 22 for full model output.

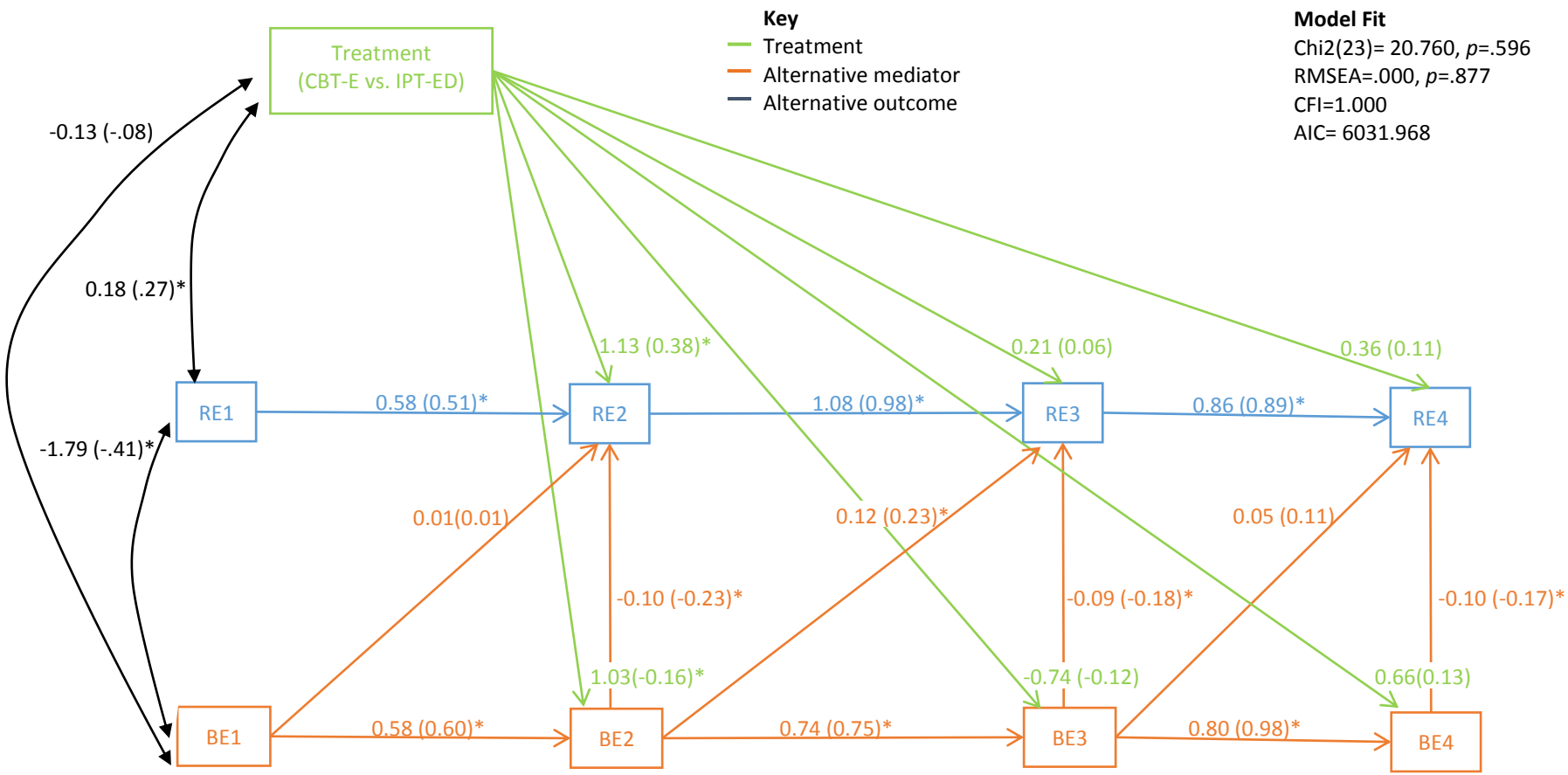


Figure 27: Alternative autoregressive structural equation model of the impact of treatment on regularity of eating (RE) as an outcome via frequency of binge eating (BE) as an alternative mediator. Numeral denotes treatment week. Controlling for shared measurement variance, depression, eating disorder diagnosis, presence of binge eating in the month prior to starting treatment, age and weight concern at baseline (standardised estimates in brackets). Significance denoted by * based on bootstrapped bias-corrected confidence intervals (1000 resamples). Missing data accounted for using full information maximum likelihood estimation.

Table 12: Comparison of the cumulative indirect and total effects of the intervention (CBT-E) on regularity of eating (RE) each week controlling for shared measurement variance, depression, eating disorder diagnosis, presence of binge eating in the month prior to starting treatment, age and weight concern at baseline

Impact of Treatment Allocation on Regularity of Eating	Treatment Week	B	SE	Bias-Corrected 95% CI of B Coefficient		β
				Lower	**Upper**	
Indirect effects	2	0.11	0.06	0.02	0.30	0.04
	3	1.36	0.25	0.94	1.95	0.42
	4	1.32	0.20	0.91	1.73	0.45
Total effects	2	1.24	0.18	0.89	1.58	0.42
	3	1.57	0.18	1.18	1.91	0.48
	4	1.68	0.18	1.35	2.05	0.54

Notes. B=unstandardized estimate, SE=standard error, CI= confidence interval, β=standardised estimate. CIs are bias corrected based on 1000 bootstraps. Indirect and total effects of treatment incorporate transmission of treatment effects via autoregressive paths (rather than each time point in isolation).

5.4.5. Study 1d: An Investigation into the Relationship between Change in the Mediator and Change in Outcome using Latent Growth Curve Modelling

To directly assess the relationship between change in regularity of eating (purported mediator) and change in frequency of binge eating (the outcome), the data were modelled using a parallel process latent growth curve model. Latent growth curve models were run separately for the purported mediator and the outcome, checking for both linear and non-linear trends. These models then formed the basis of the parallel process latent growth curve model. Change was modelled so that it equated to the mean amount of change occurring during the whole period of study (i.e. the first four weeks of treatment), rather than amount of change occurring each week of treatment in order to account for differences in rates of change between variables (see chapter 4 for more details).

Following problems with model convergence, it was decided that the factor weightings denoting the rate of change in each variable (on the slope factor) would be constrained in order to reduce the demands on the model estimation process by reducing the number of paths the model had to estimate. Factor weightings for the slope were constrained based on the factor weightings of initial, simpler models that had converged (rounded to the first decimal place), which demonstrated good initial fit. All models at this stage were the basic models (unbootstrapped, not accounting for missing data). For the purported mediator, this model included treatment allocation only as problems with convergence occurred on the inclusion of the relevant non-specific predictors. This model indicated that change in the purported mediator

(regularity of eating) was non-linear, with approximately 80% of change occurring by the second week of treatment, and 100% of change achieved by week 3. Model fit was good; Chi2(5)=6.782, *p*=.237, RMSEA=.061, *p*=.364, AIC 1143.094. Subsequent models of the purported mediator therefore had the factor weightings of the slope constrained at weeks 1-4 to be 0, .80, 1, 1. For the outcome, factor weightings were based on the basic model, examining only the growth curve as problems with convergence occurred once treatment allocation was added to the model. This model indicated that change in the outcome (frequency of weighing) was approximately linear in this variable, with 20% of change occurring by week 2 and 50% by week 3. Model fit was good; Chi2(3)=2.558, *p*=.465, RMSEA=.000, *p*=.566, AIC= 1757.853. Subsequent models therefore had the factor weightings of the slope constrained at weeks 1-4 to be 0, .20, .50, 1. The results of the final parallel process latent growth curve model are reported below.

5.4.5.1. Effect of Treatment on Change in the Purported Mediator (Regularity of Eating)

As found in studies 1a, 1b and 1c, there was a significant difference between treatments, with patients in CBT-E showing an increase in their regularity of eating by 1.36 more than patients in IPT-ED (see Figure 28). This suggests that CBT-E significantly increased the purported mediator (regularity of eating) during the first 4 weeks of treatment.

5.4.5.2. Effect of Treatment on Change in the Outcome (Frequency of Binge eating) when Accounting for the Purported Mediator (Regularity of Eating)

There was no significant difference between treatments in terms of frequency of binge eating (see Figure 28). This suggests that the outcome (frequency of binge eating) did not differ significantly in CBT-E compared to IPT-ED by the end of the first 4 weeks of treatment. This corroborates findings in studies 1b and 1c, that there were few differences between treatments at each treatment week, at least when accounting for the purported mediator.

5.4.5.3. Indirect Effect of Treatment on the Outcome (Frequency of Binge eating) via the Purported Mediator (Regularity of Eating)

There was a significant relationship between change in regularity of eating (purported mediator) and change in frequency of binge eating (outcome) during this period, with change in regularity of eating being associated with a decrease in frequency of binge eating of -1.80. This suggests that there was a significant relationship between change in the purported mediator and change in the outcome during this period.

Although there was a trend towards an indirect effect of CBT-E decreasing frequency of binge eating via regularity of eating, this was non-significant although the standardised coefficient suggested a moderate effect size (see Table 13). Confidence intervals were wide suggesting variability in the indirect effect. Close examination of the model constants and path coefficients suggested that there was a distinct pattern of relationships between regularity of eating and frequency of binge eating in CBT-E and IPT-ED, suggesting different mediators between treatments. This pattern was consistent with the idea that change in frequency of binge eating in CBT-E was

mediated by regularity of eating, at least during the first 4 weeks of treatment. See Appendix 23 for Stata code and Appendix 24 for full model output.

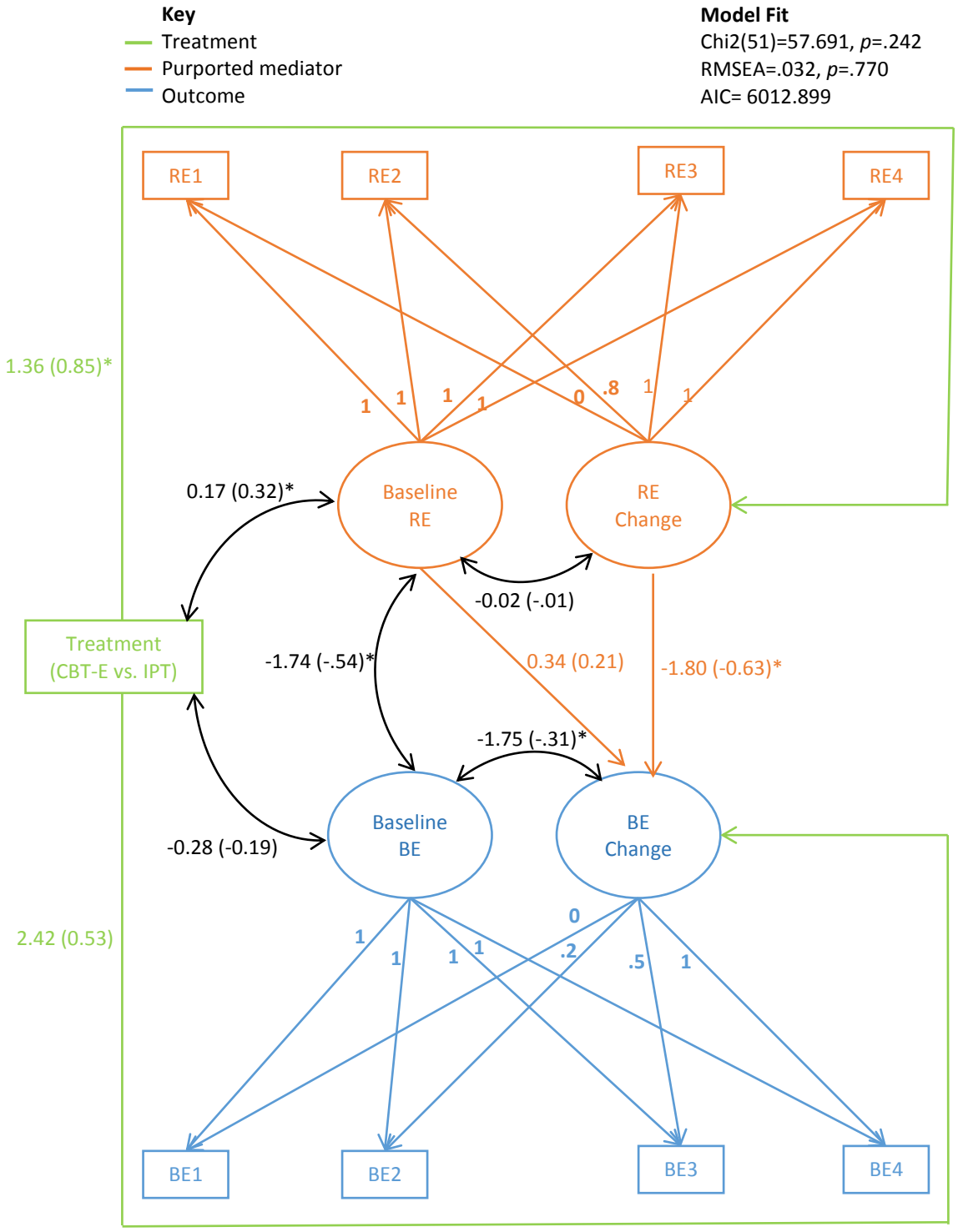


Figure 28: Parallel process latent growth curve model examining the impact of treatment allocation on frequency of binge eating (BE) via regularity of eating (RE). Numeral indicates treatment week. Controlling for depression, eating disorder diagnosis, presence of binge eating in the month prior to starting treatment, age and weight concern at baseline (standardised estimates in brackets). Significance denoted by * based on bias-corrected confidence intervals (1000 bootstrap resamples).

Table 13: Comparison of the indirect (via change in regularity of eating only) and total effects of the intervention (CBT-E) on frequency of binge eating (BE) during the first four weeks of treatment, controlling for depression, eating disorder diagnosis, presence of binge eating in the month prior to starting treatment, age and weight concern at baseline

Impact of Treatment Allocation on Frequency of Binge Eating	B	SE	Bias-Corrected 95% CI of B Coefficient		β
			Lower	**Upper**	
Indirect effects	-2.45	2.19	-7.42	0.02	-0.54
Total effects	-0.03	0.57	-1.25	0.83	-0.01

Notes. B=unstandardized estimate, SE=standard error, CI= confidence interval, β=standardised estimate. CIs are bias corrected based on 1000 bootstrap resamples. Indirect and total effects of treatment incorporate transmission of treatment effects via autoregressive paths (rather than each time point in isolation).

5.5. Summary of Results

The purported mediator (regularity of eating) increased significantly in CBT-E but changed little in IPT-ED. This was consistent across all four studies. Study 1a suggested that frequency of binge eating (outcome) decreased at different rates in CBT-E and IPT-ED, with patients in CBT-E experiencing a rapid decrease in frequency of binge eating, which slowed towards the end of the four week period, and patients in IPT-ED experiencing little change initially, followed by a decrease. Studies 1b and 1c also demonstrated different relationships between treatment, regularity of eating (purported mediator) and frequency of binge eating (outcome) at different time points. When accounting for the purported mediator (regularity of eating), which was significantly related to the outcome (frequency of binge eating) there was no longer a significant relationship between treatment allocation and outcome. These findings were consistent across studies 1a, 1b and 1d. Reversing the relationship in study 1c did not suggest that change in regularity of eating (the purported mediator) was better explained by changes in the outcome (frequency of binge eating). There was evidence of a significant indirect effect of CBT-E on the outcome (frequency of binge eating) via regularity of eating when examining change at discreet treatment weeks in study 1b, and a non-significant trend towards this for change over the four week period as a whole in study 1d.

Overall, the results support regularity of eating as a mediator of the effect of CBT-E on the outcome (frequency of binge eating). However, as temporal precedence could not be established, the results are not definitive.

Chapter 6. Study 2: Frequency of Weighing as a Mediator of the Effect of Enhanced Cognitive Behavioural Therapy on Concern about Weight

6.1. Background

As discussed in chapters 1 and 2, patients with eating disorders demonstrate high levels of concern about weight, and many engage in frequent weighing as a way of monitoring weight change. In enhanced cognitive behaviour therapy (CBT-E) a core intervention early in stage 1 of treatment is the "weekly weighing" procedure, which is implemented in the first week of treatment after session 0 (see Chapter 3 for treatment timings). This intervention has several components. First, patients are educated about body weight and the correct interpretation of body weight fluctuations. Weekly in-session weighing is then used as an opportunity to educate the patient and test his or her fears surrounding weight to help the patient learn to correctly interpret weight (i.e. the number on the scale) and 'weight change' over time. Interpretation is done collaboratively, with the patient and therapist agreeing the number on the scale, and discussing its interpretation. Patients are taught never to interpret weight based on a single reading, but rather to look at weight trends over a 4 week period. They are also encouraged to avoid weighing themselves outside of sessions to avoid the patient becoming focussed on weight fluctuations rather than the overall trend.

6.2. Aim / Hypothesis

The cognitive behavioural model posits that frequent weighing highlights inconsequential fluctuations in weight (typically due to changes in hydration), reinforcing the patients concern about weight gain, and leading them to further restrict their diet (Fairburn, Cooper, Shafran, et al., 2008b). It therefore hypothesises that by increasing patients' understanding of body weight and weight change and helping them reduce their frequency of weighing, this will start to reduce the patient's over-evaluation of weight (see Figure 29).

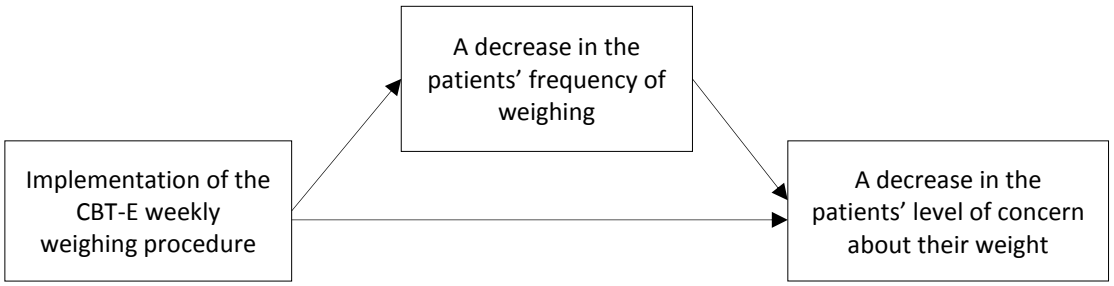


Figure 29: Illustration of the purported mediational relationship between enhanced cognitive behaviour therapy (CBT-E), frequency of weighing and concern about weight

In contrast, interpersonal psychotherapy for eating disorders (IPT-ED) does not directly address or refer to weighing frequency. Whilst concern about weight might decrease in IPT-ED via increases in self-esteem related to improved interpersonal functioning, such change would likely occur more slowly than in CBT-E (Murphy et al., 2009, 2012).

The aim of this study was to assess whether change in frequency of weighing mediates change in concern about weight in CBT-E compared to IPT-ED. It was hypothesised that patients in CBT-E would decrease their frequency of weighing more rapidly than in IPT-ED and that this decrease would mediate change in concern about weight in CBT-E. It was also hypothesised that the effect of this intervention might differ depending on

initial levels of frequency of weighing, therefore this was tested as a moderator (see Analytic considerations for more details).

6.3. Method

6.3.1. Experimental Procedure

Patients in CBT-E were weighed once a week by their therapist as per the treatment protocol. In IPT-ED patients were not weighed. Patients in both treatments completed the 'Eating Questionnaire (EQ)' on a weekly basis immediately prior to the treatment session (see chapter 3 for more detail). For the purposes of this study a period of six weeks was selected to account for change in both the purported mediator (frequency of weighing) and the outcome (concern about weight). Change in frequency of weighing was hypothesised to be rapid. In contrast, change in the outcome (concern about weight) was hypothesised to be slower; firstly due to the fact that cognitive change is hypothesised to be slower, and secondly, interpretation of weight was meant to be based on trends over a four week period. A longer period was not chosen to prevent confounding with later interventions carried out in stage 3 of CBT-E that might have an impact on these variables concern about weight (e.g. addressing dietary rules and introducing previously avoided foods back into one's diet).

6.3.2. Study Measures

6.3.2.1. Frequency of Weighing

Frequency of weighing was assessed using item 16 of the EQ (see Appendix 6). This item was a count of the number of times the patient had weighed his/herself outside

of the treatment session during the past week. To account for the fact that patients in CBT-E were weighed once a week in session as standard, frequency of weighing was reduced by 1 for patients in IPT-ED to make frequency of weighing equivalent in the two treatments.

6.3.2.2. Concern about Weight

Concern about weight was assessed using item 26 on the EQ (see Appendix 6). This item assessed how worried patients had been about their weight during the past week on a Likert scale of 0-6, where 0 indicated 'not at all' and 6 indicated 'markedly'.

6.3.3. Analytic considerations

6.3.3.1. Potential Confounders

Potential unmeasured confounders were; change in body mass index (BMI) and exposure to weight salient cues. Neither of these variables were captured in the research data. However, research suggests that BMI changes little during treatment for non-underweight patients (Fairburn et al., 2009, 2015).

6.3.3.2. Frequent Weigher Status as a Potential Moderator

Baseline frequency of weighing was tested as an additional moderator in this study and was based on the frequency of weighing reported in the Eating Disorder Examination undertaken in the main efficacy trial prior to commencing treatment (see Chapter 4). A binary variable was created to identify frequent and non-frequent

weighers. Frequent weighers were defined as those patients weighing themselves more than 4 times a month, equivalent to more than once per week.[5]

In study 2a (A Preliminary Investigation using Multilevel Modelling for Statistical Mediation), an interaction was included between treatment allocation and frequent weigher status to assess whether the impact of being a frequent weigher on frequency of weighing (purported mediator) and concern about weight (outcome) differed by treatment. Subsequently a three way interaction between treatment allocation, frequent weigher status and treatment week to assess whether there were differences between treatments over time between frequent and non-frequent weighers. In study 2b (An In-Depth Analysis using Autoregressive Structural Equation Modelling for Statistical Mediation), frequent weigher status was added as a simple interaction term between treatment and frequent weigher status at each time point.

6.4. Results

6.4.1. Non-Normality and Missing Data

Examination of the data suggested non-normality, particularly in the frequency of weighing variable, which was positively skewed throughout treatment, suggesting low levels of frequent weighing. In contrast, concern about weight was negatively skewed at baseline, suggesting high levels of concern about weight. Missing data was mostly due to attrition.

[5] Whilst weighing avoiders (patients who were deliberately avoiding weighing themselves) might also be expected to respond differently to the weekly weighing procedure, few patients identified as avoiding weighing themselves at baseline [N=11], and by the start of treatment this number had reduced further [N=9]. Therefore, it was not feasible to test weighing avoiders as a separate category.

6.4.2. Study 2a: A Preliminary Investigation using Multilevel Modelling for Statistical Mediation

6.4.2.1. Step 1: Impact of the Intervention (CBT-E) on the Purported Mediator (Frequency of Weighing)

Frequency of weighing (purported mediator) differed between frequent and non-frequent weighers in CBT-E and IPT-ED (see Figure 30 and Figure 31). There was a significant three-way interaction between treatment allocation, baseline frequent weigher status and treatment week. Frequency of weighing decreased rapidly in frequent weighers in CBT-E at a rate of -3.25 (p<.001) more each week compared to non-frequent weighers and frequent weighers in IPT-ED (see graph 1, Figure 32). This was qualified by a significant quadratic trend of 0.41 (p=.004) each week suggesting that rate of change plateaued by week 3. In contrast to frequent weighers in CBT-E, frequency of weighing did not change significantly in frequent weighers in IPT-ED, nor in non-frequent weighers in either treatment. See Appendix 25 for Stata code and Appendix 26 for full model output.

6.4.2.2. Step 2: Impact of the Intervention (CBT-E) on the Outcome (Concern about Weight)

Concern about weight (outcome) decreased at a similar rate in frequent and non-frequent weighers in CBT-E and IPT-ED (see Figure 30 and Figure 31). The interaction between treatment allocation and treatment week was not significant (B=0.04, p=.434), suggesting similar rates of change in the two treatments. Overall, concern about weight decreased by -0.36 (p<.001) each week (see graph 2, Figure 32). This was qualified by a significant quadratic trend of 0.03 (p=.015) suggesting that rate of

change slowed by week 6 of treatment. Baseline frequent weigher status did not significantly moderate the effect of treatment and its interactive effect with treatment was therefore dropped from the final model. See Appendix 25 for Stata code and Appendix 27 for full model output.

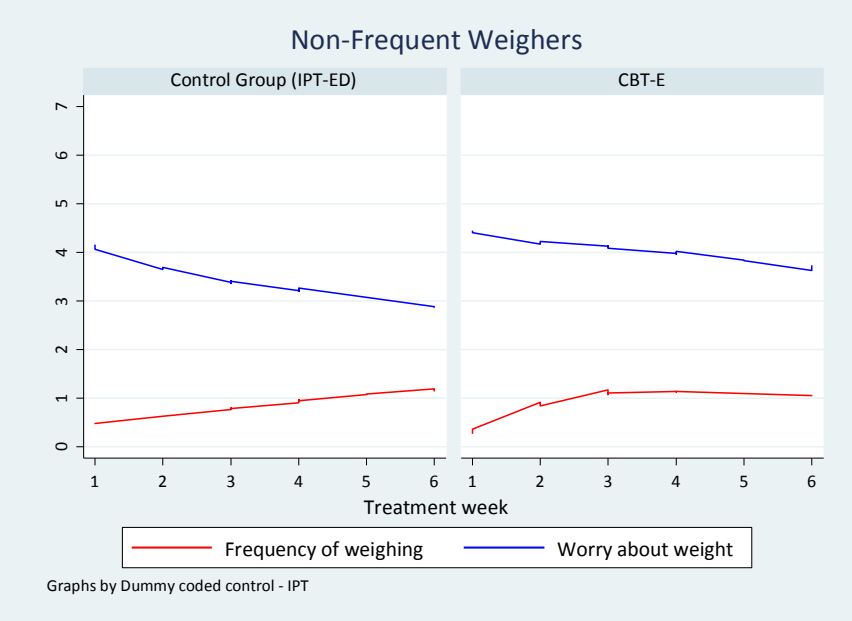


Figure 30: Line graph plotting the observed means (smoothed curve) of frequency of weighing and concern about weight in non-frequent weighers in the IPT-ED and CBT-E

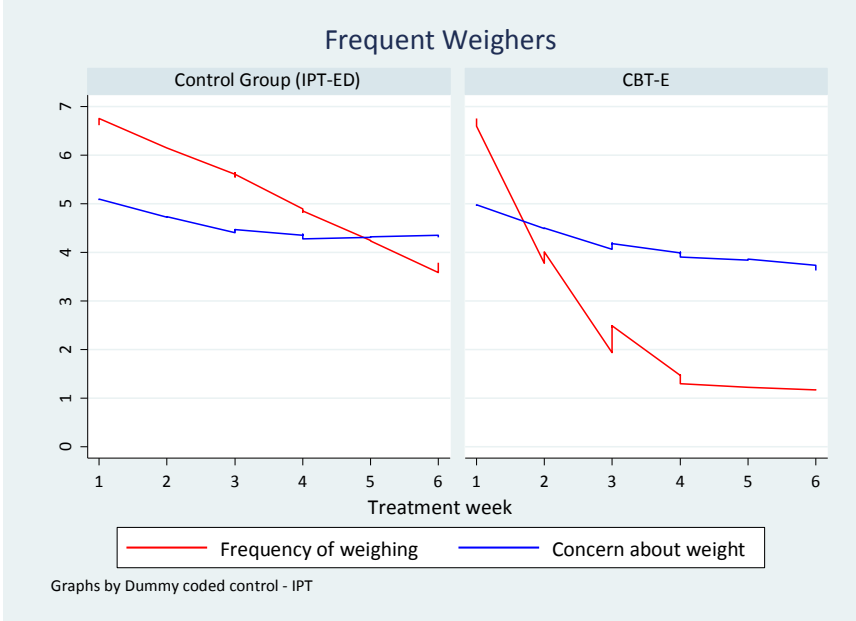


Figure 31: Line graph plotting the observed means (smoothed curve) of frequency of weighing and concern about weight in frequent weighers in the IPT-ED and CBT-E

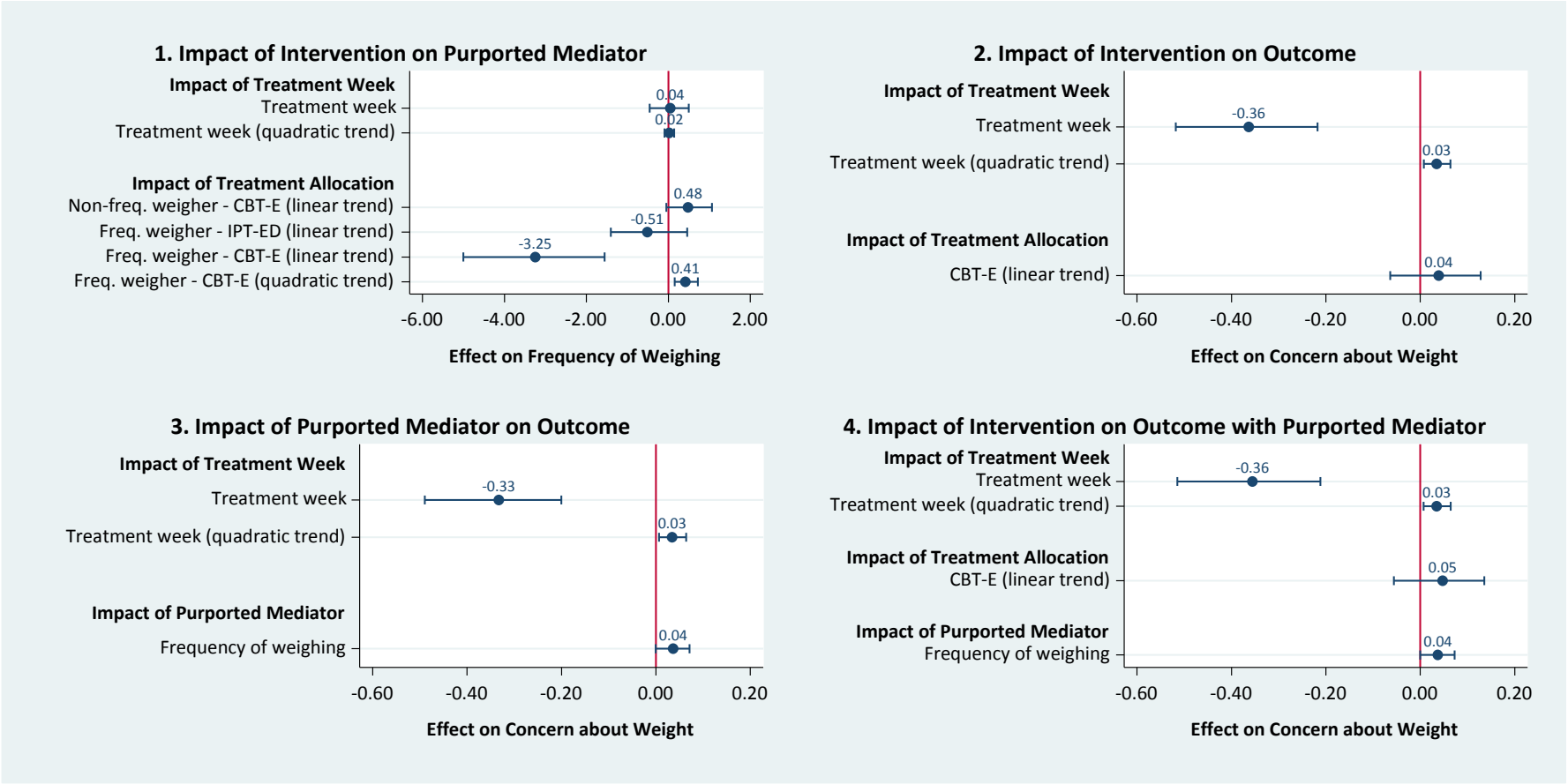


Figure 32: Coefficient plots with 95% bias-corrected confidence intervals (1000 bootstrap resamples) comparing the effect of CBT-E on frequency of weighing and concern about weight when frequency of weighing was / was not included as a mediator. Impact of treatment allocation represents an interaction between treatment allocation and treatment week. Controlling for the effects of baseline frequent weigher status, self-esteem and weight concern at baseline.

6.4.2.3. Step 3: Relationship between the Purported Mediator (Frequency of Weighing) and the Outcome (Concern about Weight)

Frequency of weighing (purported mediator) was positively associated with concern about weight (outcome), with a one point increase in frequency of weighing being associated with a 0.04 (*p*=.047) increase in concern about weight (see graph 3, Figure 32). However, there was no clear temporal precedence for change in these variables and change in one variable appeared to have little observable relationship to change in the other (see Figure 30 and Figure 31). Whilst concern about weight decreased in both frequent and non-frequent weighers alike, this appeared to be irrespective of change in frequency of weighing. See Appendix 25 for Stata code and Appendix 28 for full model output.

6.4.2.4. Step 4: Impact of the intervention (CBT-E) on the Outcome (Concern about Weight) when Accounting for the Purported Mediator (Frequency of Weighing)

When including treatment allocation and frequency of weighing (purported mediator) in the same model, there was little change in the impact of treatment on concern about weight, with this relationship remaining non-significant (B=0.05, *p*=.337) despite frequency of weighing being significantly related to concern about weight (see graph 4, Figure 32). This does not support frequency of weighing being a mediator of the effect of CBT-E on concern about weight during treatment. See Appendix 25 for Stata code and Appendix 29 for full model output.

Lagged models assessing the effect that prior level of the purported mediator had on subsequent outcome found little evidence of a relationship between earlier measures

of frequency of weighing (purported mediator) and the outcome (concern about weight).

6.4.3. Study 2b: An In-Depth Analysis using Autoregressive Structural Equation Modelling for Statistical Mediation

To verify the results of study 2a, namely that there was no indirect relationship between treatment and concern about weight via frequency of weighing in CBT-E, an autoregressive structural equation model was used to directly model the relationships between treatment, frequency of weighing and concern about weight. Baseline frequent weigher status was included as a significant moderator of the effect of treatment on frequency of weighing, but not concern about weight, as determined in the preliminary analyses using multilevel modelling in study 2a.

6.4.3.1. Effect of the intervention (CBT-E) on the Purported Mediator (Frequency of Weighing)

There was a significant interaction between treatment allocation and baseline frequent weigher status, with frequency of weighing (purported mediator) being significantly lower in frequent weighers in CBT-E compared to non-frequent weighers and frequent weighers in IPT-ED by -3.73 in week 2 and -2.36 in week 3 (see Figure 33). There was also a significant difference between non-frequent weighers in CBT-E compared to IPT-ED, with non-frequent weighers in CBT-E weighing themselves by -0.79 less compared to patients in IPT-ED in week 4 of treatment. However, the strength of relationship of prior frequency of weighing to frequency of weighing at later time points did not vary significantly during this period; Chi2(4)=8.56, *p*=.073. This suggests that CBT-E significantly reduced frequency of weighing in frequent

weighers, and either reduced or prevented increases in frequency of weighing in non-frequent weighers i.e. treatment had a significant effect on the purported mediator during this period.

6.4.3.2. Effect of the intervention (CBT-E) on the Outcome (Concern about Weight) when Accounting for the Purported Mediator (Frequency of Weighing)

There were no significant differences in concern about weight (outcome) between treatments at any week when accounting for the purported mediator (see Figure 33). Furthermore, the strength of relationship between prior concern about weight to concern about weight at later time points did not vary significantly during this period; Chi2(4)=7.69, *p*=.104. This suggests that treatment allocation did not affect concern about weight i.e. treatment did not have a significant effect on the outcome when the purported mediator was accounted for.

6.4.3.3. Indirect Effect of the Intervention (CBT-E) on the Outcome (Concern about Weight) via the Purported Mediator (Frequency of Weighing)

Although there was a weak correlation between frequency of weighing (purported mediator) and concern about weight (outcome) at baseline, there was no significant relationship between these two variables during treatment, except at week 4, where the effects of the previous week and the current week combined to cancel each other out (see Figure 33). There was no evidence for an indirect effect of CBT-E decreasing concern about weight via frequency of weighing in either frequent or non-frequent weighers (see Table 14). This suggests that frequency of weighing was unrelated to change in concern about weight in CBT-E i.e. the purported mediator did not mediate

the effect of CBT-E on the outcome. See Appendix 30 for Stata code and Appendix 31 for full model output.

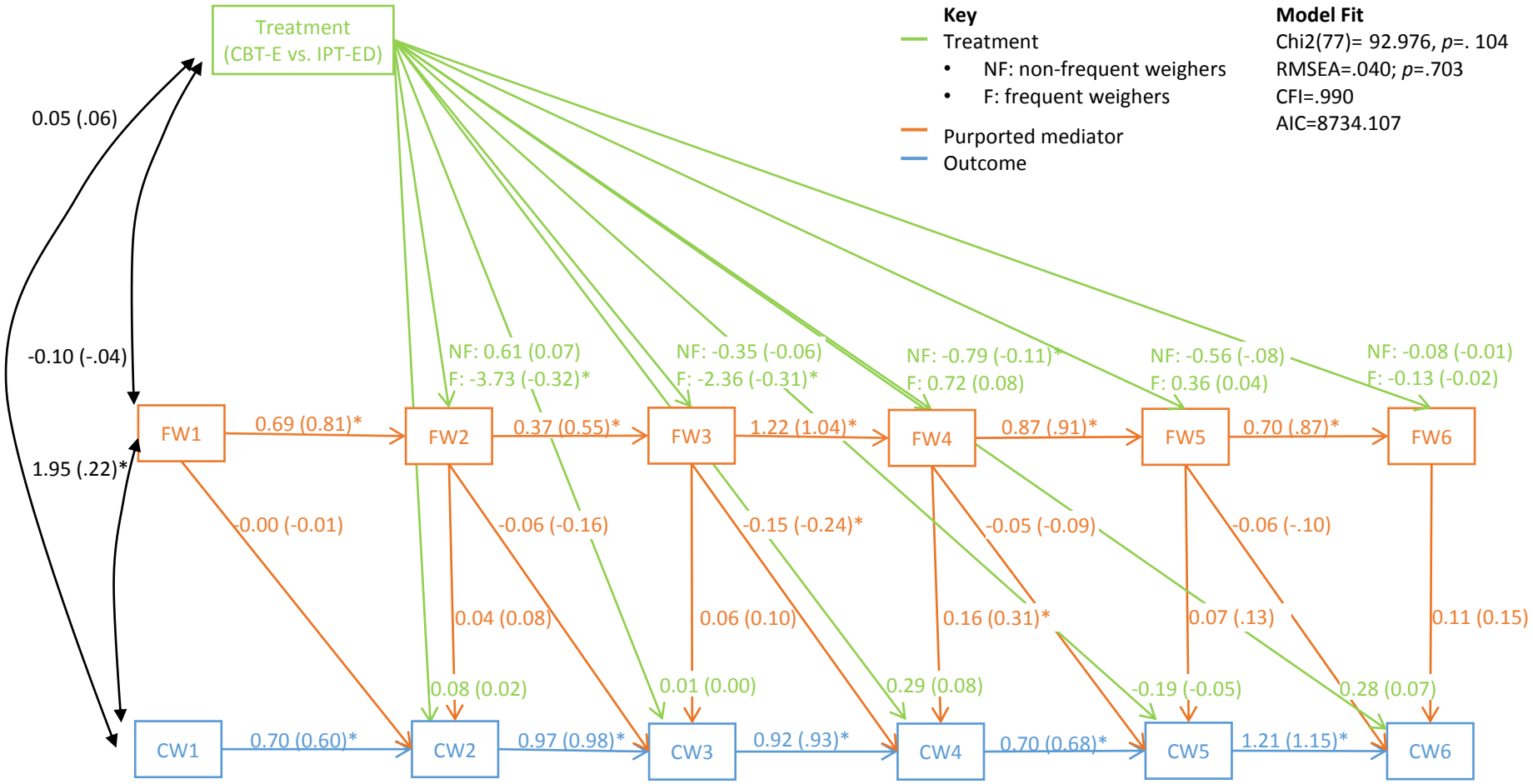


Figure 33: Autoregressive structural equation model of the impact of treatment on concern about weight (CS) via frequency of weighing (FW) in frequent and non-frequent weighers. Numeral indicates treatment week. Controlling for shared measurement variance, self-esteem and weight concern at baseline (standardised estimates in brackets). Effect of treatment allocation and baseline frequent weigher status on frequency of weighing modelled as an interaction. Significance denoted by * based on bootstrapped bias-corrected confidence intervals (1000 resamples). Missing data accounted for using full information maximum likelihood estimation.

Table 14: Comparison of the cumulative indirect and total effects of the intervention (CBT-E) on concern about weight each week in non-frequent weighers, controlling for shared measurement variance, self-esteem and weight concern at baseline

Impact of Treatment Allocation on Concern about Weight	Treatment Week	Non-Frequent Weighers					Frequent Weighers				
		B	**SE**	**Bias-Corrected 95% CI of B Coefficient**		**β**	**B**	**SE**	**Bias-Corrected 95% CI of B Coefficient**		**β**
				Lower	**Upper**				**Lower**	**Upper**	
Indirect effects	2	0.02	0.05	-0.02	0.18	0.01	-0.13	0.14	-0.45	0.10	0.10
	3	0.05	0.20	-0.33	0.45	0.01	-0.12	0.18	-0.50	0.22	0.22
	4	-0.08	0.24	-0.56	0.37	-0.02	-0.18	0.19	-0.61	0.14	0.14
	5	0.09	0.23	-0.40	0.48	0.03	-0.16	0.18	-0.61	0.15	0.15
	6	0.02	0.05	-0.02	0.18	-0.04	-0.27	0.23	-0.71	0.17	0.17
Total effects	2	0.10	0.21	-0.31	0.51	0.03	-	-	-	-	-
	3	0.06	0.25	-0.45	0.53	0.02	-	-	-	-	-
	4	0.21	0.28	-0.30	0.76	0.06	-	-	-	-	-
	5	-0.10	0.30	-0.72	0.45	-0.03	-	-	-	-	-
	6	0.10	0.21	-0.31	0.51	0.03	-	-	-	-	-

Notes. B=unstandardized estimate, SE=standard error, CI= confidence interval, β=standardised estimate. CIs are bias-corrected based on 1000 bootstraps. Indirect and total effects of treatment incorporate transmission of treatment effects via autoregressive paths (rather than each time point in isolation). Indirect and total effects are equivalent in frequent weighers as there was no direct path to the outcome in this group given that baseline frequent weigher status was not a moderator of the outcome.

6.5. Summary of Results

The purported mediator (frequency of weighing) decreased significantly in frequent weighers in CBT-E, but changed little in non-frequent weighers and frequent weighers in IPT-ED. Study 2a suggested that concern about weight (outcome) decreased at a similar rate in both treatments, and there were no differences between treatments or by frequent weigher status. Accounting for the purported mediator (frequency of weighing) did not affect the impact of treatment allocation on the outcome (concern about weight), remaining non-significant in both studies. Whilst study 2a suggested a weak positive relationship between the purported mediator (frequency of weighing) and the outcome (concern about weight), study 2b suggested that this relationship was present at the start of treatment only, and that there was little evidence for an ongoing relationship between the two variables during treatment.

Overall, there was no evidence that the purported mediator (frequency of weighing) accounted for the effect of CBT-E on the outcome (concern about weight), despite CBT-E's significant impact on frequency of weighing early on in treatment.

Chapter 7. Study 3: Frequency of Shape Checking as a Mediator of the Effect of Enhanced Cognitive Behaviour Therapy on Concern about Shape

7.1. Background

A key procedure that occurs later in enhanced cognitive behaviour therapy (CBT-E) is the "shape checking" intervention, which is implemented as part of the shape concern module in stage three (weeks 5-14) of CBT-E (see chapters 1 and 2). This procedure focuses on helping patients reduce frequent and unhelpful forms of shape checking behaviours e.g. body touching / measuring, and excessive mirror use. The intervention has two key stages; firstly, a 'monitoring' stage in which the patient is asked to observe and record their shape checking behaviours for two days, one a work day and the other a day off work, and secondly, an 'addressing' stage in which the patient is supported to consider the utility of these checking behaviours and reduce their frequency. The timing for the implementation of this intervention is personalised to the patient (see chapter 3). Once started, this procedure remains an active agenda item until the end of treatment.

7.2. Aim/Hypothesis

The cognitive behavioural theory of the maintenance of eating disorder psychopathology proposes that frequent shape checking reinforces over-evaluation of shape and weight by providing misleading information about body shape, making it seem larger than it really is. It therefore hypothesises that by helping patients reduce and reinterpret their shape checking, this in turn will start to reduce their over-evaluation of shape. It was anticipated that in CBT-E reductions in concern about shape would be specifically related to the implementation of the shape checking intervention via this process (see Figure 34).

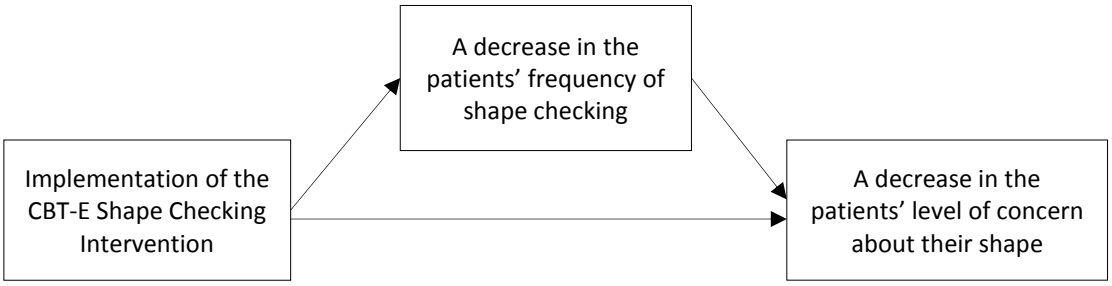


Figure 34: Illustration of the purported mediational relationship between enhanced cognitive behaviour therapy (CBT-E), frequency of shape checking and concern about shape

In comparison, interpersonal psychotherapy for eating disorders (IPT-ED) does not directly address or refer to shape checking. Whilst concern about shape might decrease in IPT-ED via increases in self-esteem related to improved interpersonal functioning, such change would likely occur more slowly than in CBT-E (Murphy et al., 2009, 2012).

The aim of this study was to assess whether change in patients' frequency of shape checking mediates change in their concern about shape to a greater extent in CBT-E compared to IPT-ED. It was hypothesised that patients in CBT-E would decrease their

frequency of shape checking more rapidly than in IPT-ED and that this decrease would mediate change in concern about weight in CBT-E. It was hypothesised that this change would be related to the implementation of the shape checking intervention.

7.3. Method

7.3.1. Experimental Procedure

Patients completed the 'Eating Questionnaire' (EQ) on a weekly basis immediately prior to the treatment session (with the exception of stage four when sessions occurred fortnightly, see chapter 3 for more detail). After each session trial therapists recorded which interventions had been a major focus of treatment in order to capture when the personalised interventions in CBT-E were implemented for each patient using the 'Intervention Record Sheet (IRS)'. As part of this, both implementation of shape checking monitoring and shape checking addressing were recorded.

7.3.2. Study Measures

7.3.2.1. Implementation of the Shape Checking Intervention

Implementation of the shape checking intervention was a binary variable indicating whether the intervention had been implemented or not in that session, based on the IRS completed by the trial therapist (see Appendix 10 and Appendix 11)

7.3.2.2. Frequency of Shape Checking

Frequency of shape checking was assessed using items 18 and 19 of the EQ (see Appendix 6). These items rated frequency of mirror checking and body checking (touching / measuring) during the past week using a Likert scale of 1-5, where 1

indicated no checking and 5 indicated a frequency of many times per day. Mirror and body checking were moderately-strongly positively correlated (Pearson *r* = .56-.82 at each time point).

7.3.2.3. Concern about shape

Concern about shape was assessed using item 27 on the EQ (Appendix 6). This item assessed how worried patients had been about their shape during the past week on a Likert scale of 0-6, where 0 indicated 'not at all' and 6 indicated 'markedly'. Whilst other items corresponding to the shape concern subscale of the validated 'eating disorder examination' questionnaire were available, because several items in this subscale would not have been relevant for the shape checking intervention (e.g. item 28 on the EQ, which corresponded to body avoidance) it was decided to use a representative single item as a measure. Item 27 was selected as best representing concern about body shape.

7.3.3. Analytic considerations

7.3.3.1. Potential Confounders

Potential unmeasured confounders were change in shape (possibly due to change in BMI) and exposure to shape salient cues. Neither of these variables were captured in the research data. However, as discussed in study 2, research suggests that BMI changes little during treatment for non-underweight patients (Fairburn et al., 2009, 2015).

7.3.3.2. Incorporating Personalisation of Treatment

In comparison to the other studies in this thesis, which used a fixed variable denoting treatment allocation, the implementation of the shape checking intervention in CBT-E varied between patients and over time. This analysis therefore focused on the differences between patients who received the shape checking intervention and those who did not. An important consideration was therefore how best to incorporate the personalised implementation of the intervention into the analyses. There were a number of ways the variable could have been coded, each representing different ways of conceptualising the impact of the intervention. Three alternative coding methods, along with their implications, are described below.

Switch on/off method. In this coding, 0 indicates that the intervention was not used in a session (i.e. it is inactive), and 1 indicates that it was used in a session (i.e. it is active) (see Table 15). This coding makes the assumption that the intervention is only active when the intervention is a major treatment focus, and that these effects would only be seen immediately after (i.e. the effects are impermanent).

Switch on method. In this method, 0 indicates that the intervention has not been implemented (i.e. it is inactive), and 1 indicates that it has been implemented (i.e. it is active) (see Table 15). Unlike the switch on/off method, this coding assumes an ongoing effect following the initial implementation of the intervention. This could be either because the intervention itself affects a 'permanent' or 'semi-permanent' change in behaviour or that the intervention, or, although not a major session focus, the intervention continues to be a small part of the overall treatment.

Total dose method. For this method of coding, 0 indicates the intervention is inactive, 1 indicates the first session that the intervention was implemented and each subsequent increase in number denotes receipt of another session in which the intervention was a major treatment focus (see Table 15). This assumes that each additional session has an additive effect on the patient's outcome.

Table 15: Examples of coding schemes for personalisation of treatment variable

Treatment Week	Was the intervention a major focus in the session?	Alternative Strategies for Coding Personalisation of Treatment in the Analyses		
		Switch on/off	**Switch on**	**Total dose**
4	No	0	0	0
5	Yes	1	1	1
6	Yes	1	1	2
7	Yes	1	1	3
8	No	0	1	3
9	Yes	1	1	4
10	No	0	1	4

A switch on method of coding was selected for several reasons; firstly, the effects of CBT-E are well documented, with changes following treatment being maintained until at least 12 months follow-up (Fairburn et al., 2015). Secondly, once the shape checking intervention is implemented it continues to be an active session item throughout the remainder of treatment (Fairburn, Cooper, Shafran, et al., 2008b), even after it is no longer a major focus of treatment. Lastly, the total dose approach makes an assumption regarding the 'dose' of therapy (i.e. that therapy sessions are 'additive). There is too little information about this issue; the switch on method makes fewer assumptions in comparison.

7.3.3.3. Dealing with multiple timelines

There were two timelines within this study; treatment week, denoting time since treatment (CBT-E vs. IPT-ED) was started, and intervention week, denoting time since

the shape checking intervention was implemented in each patient in CBT-E who received this intervention.

In study 3a (A Preliminary Investigation using Multilevel Modelling for Statistical Mediation), a 'piecewise' model was fit whereby intervention week was modelled as an additional 'timeline' within the whole of the treatment. Intervention week was coded as zero until the first week of implementation of the shape checking intervention, after which it increased incrementally each week until the end of treatment. This enabled examination of the effect of implementation of the shape checking intervention from the point at which it was implemented. Because of this additional timeline, random slopes were fitted for both treatment week (as in the other studies in this thesis) and for intervention week, with intervention week nested within treatment week.

In study 3b (An In-Depth Analysis using Autoregressive Structural Equation Modelling for Statistical Mediation), a single, shared timeline was created to permit direct comparison between patients who did receive the shape checking intervention (intervention group) and those who did not (control group). In the intervention group, the purported mediator and the outcome were centred at the week the shape checking intervention was implemented in each patient. In the control group, scores were centred at the average week of implementation of the shape checking intervention in the intervention group (week 8). The analysis focused on week 0 to week 6 of the shared timeline, covering a period from just before the implementation of the shape checking intervention (week 0), the period in which patients monitored their shape checking but were not instructed to change it (week 1) and a period of five

weeks when patients would have been attempting to address their shape checking (weeks two to six). This covered a similar timeframe as study 2b, which examined the relationship of similar constructs (frequency of weighing and worry about weight).

7.3.3.4. Randomisation of Treatment and Potential Confounders

As a result of treatment personalisation, some patients in CBT-E did not receive the shape checking intervention. The control group therefore included patients in IPT-ED as well as those patients in CBT-E who did not receive the shape checking intervention. This meant that randomisation was broken in this study (see Figure 35). This was controlled for by including baseline covariates.

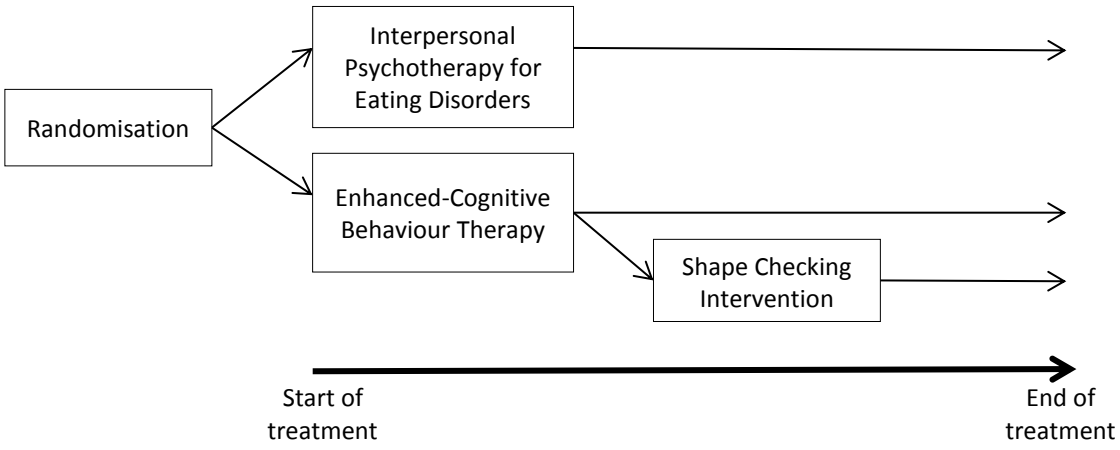


Figure 35: Illustration of implementation of shape checking intervention and the impact this has on randomisation

Treatment allocation was also a possible confounder of the effect of the shape checking intervention on the purported mediator (frequency of checking) and the outcome (concern about shape). In study 3a, treatment allocation and its interaction with treatment week were included in the model to help separate out the distinct effects of the shape checking intervention and treatment allocation. This was possible as intervention status was modelled as a piecewise trend that occurred within the 20

week period of treatment. In study 3b, this was not possible as intervention status was modelled as a binary variable. As only patients in CBT-E received the shape checking intervention, treatment allocation and intervention status had a shared membership of 82% (N=53). Of the few patients in CBT-E (N=12) who did not receive the shape checking intervention, the majority (N=8) dropped out of treatment by week 8 (week 0 of the shared timeline). This meant that the ability to separate out the effects of treatment allocation and intervention status were limited. Due to the restricted timeline, baseline levels of frequency of shape checking (purported mediator) and concern about shape (outcome) were also included as additional covariates.

A sensitivity analysis was run at the end of the analyses excluding patients from CBT-E who did not receive the shape checking intervention. No notable differences were observed.

7.4. Results

7.4.1. Non-Normality and Missing Data

Examination of the data suggested non-normality, particularly in the concern about shape variable, which was negatively skewed at baseline, suggesting high levels of concern about shape in the sample. Frequency of checking was also negatively skewed, suggesting high levels of shape checking in the sample.

7.4.2. Implementation of the Shape Checking Intervention

In total, 81.5% (N = 53) of patients allocated to CBT-E started the shape checking intervention. Of these, 98.1% (N = 52) proceeded on to address their shape checking.

On average, shape checking monitoring was implemented in week 8 of treatment (M = 7.66, SD = 2.01), remaining a key focus of treatment for 1 session (M = 1.09, SD = 0.30). Shape checking addressing started in week 9 (M = 8.85, SD = 2.02) and remained a key focus of treatment for 4 sessions (M = 4.35, SD = 2.38). For 92.5% (N = 49) of patients this intervention was an active focus of treatment for 10 weeks.

7.4.3. Study 3a: A Preliminary Investigation using Multilevel Modelling for Statistical Mediation

7.4.3.1. Step 1: Impact of the Intervention (Shape Checking Intervention) on the Purported Mediator (Frequency of Shape Checking)

Frequency of shape checking (purported mediator) changed little in the control group but decreased in the intervention group (see Figure 36). Whilst the interaction between treatment allocation and treatment week was not significant (B=-0.05, *p*=.112), this was qualified by a significant quadratic trend of 0.01 (*p*=.021), which suggested that frequency of checking increased over time in CBT-E. In contrast to CBT-E, in IPT-ED the frequency of shape checking changed little during this period. Controlling for the interaction between treatment allocation and treatment week, patients in the shape checking intervention group decreased their frequency of shape checking by -0.18 (*p*<.001) points each week of the intervention following the implementation of shape checking intervention (see Figure 37 and graph 1 in Figure 38). See Appendix 32 for Stata code and Appendix 33 for full model output.

7.4.3.2. Step 2: Impact of the Shape checking intervention on the Outcome (Concern about shape)

Concern about shape (outcome) decreased in both the control group and the intervention group (see Figure 36 and Figure 37). The interaction between treatment allocation and treatment week was not significant, and was therefore dropped from the final model. Concern about shape decreased by -0.13 (p<.001) each week of treatment. This was qualified by a significant positive quadratic trend of (B=0.004, p=.007), suggesting that rate of change slowed towards the end of treatment. In contrast, in the shape checking intervention group, concern about weight decreased by an additional -0.09 (p=.002) each week of the intervention following implementation of the shape checking intervention (see graph 2 in Figure 38). See Appendix 32 for Stata code and Appendix 34 for full model output.

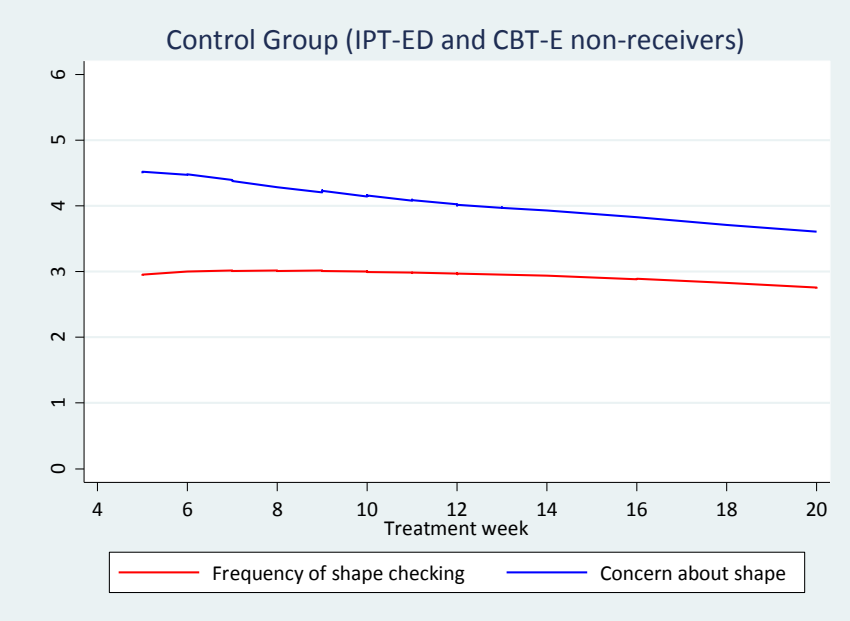


Figure 36: Line graph plotting the observed means (smoothed curve) of frequency of shape checking and concern about shape in the control group in weeks 4 to 20 of treatment

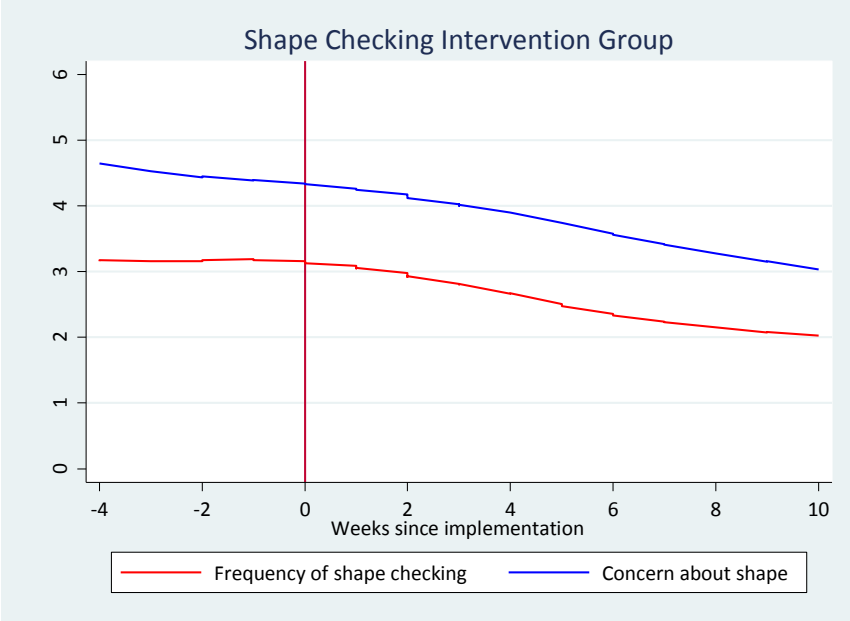


Figure 37: Line graph plotting the observed means (smoothed curve) of frequency of shape checking and concern about shape in the intervention group from 4 weeks before the shape checking intervention was implemented until 10 weeks after

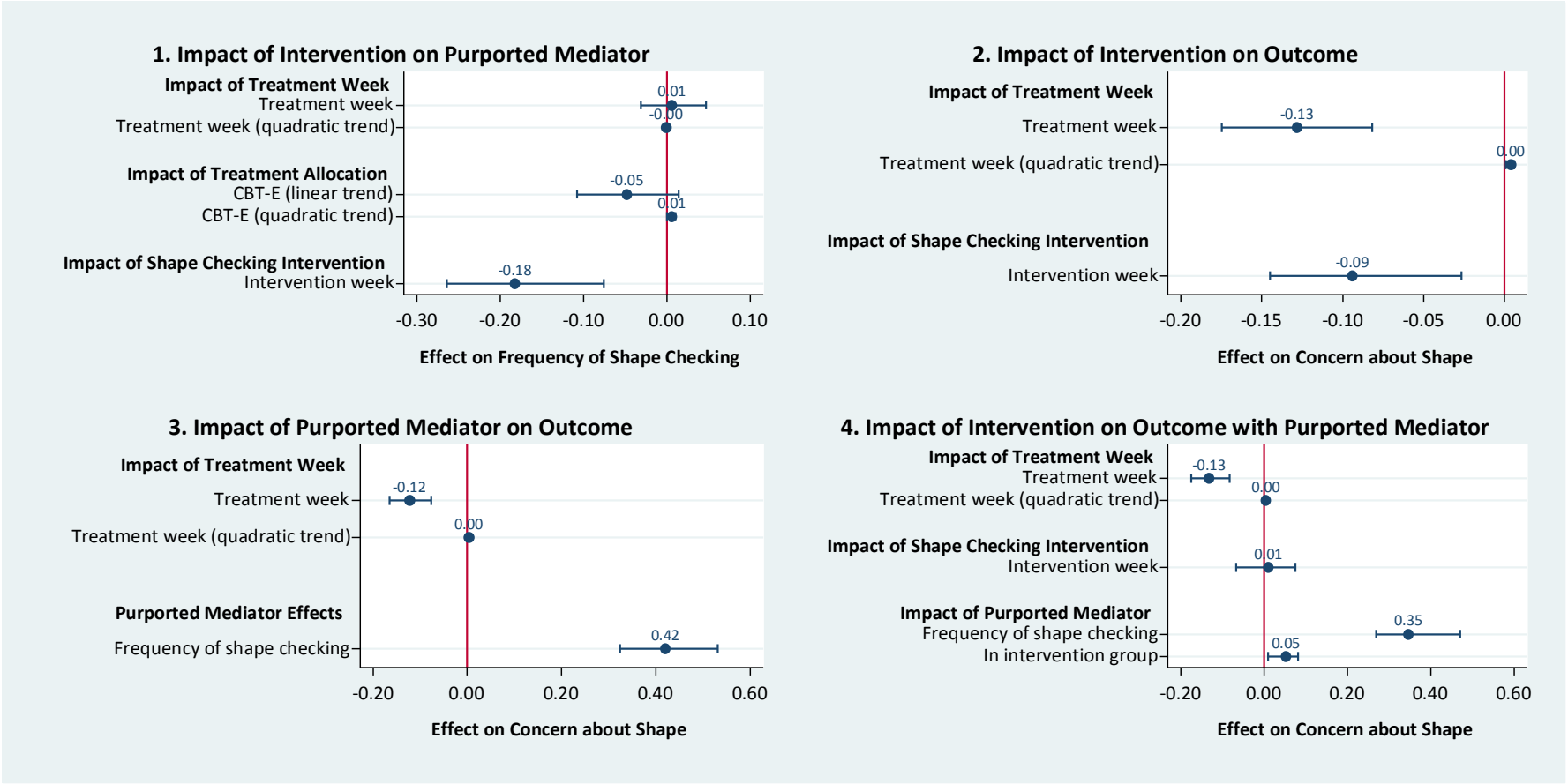


Figure 38: Coefficient plots with 95% bias-corrected confidence intervals (1000 bootstrap resamples) comparing the effect of the shape checking intervention on frequency of shape checking and concern about shape when frequency of shape checking was / was not included as a mediator. Impact of treatment allocation represents an interaction between treatment allocation and treatment week. Controlling for the effects of eating disorder duration, age, weight concern, self-esteem, depression and importance of shape at baseline.

7.4.3.3. Step 3: Relationship between the Purported Mediator (Frequency of Shape Checking) and the Outcome (Concern about shape)

Frequency of shape checking (purported mediator) was positively associated with concern about shape, with a one point increase in frequency of shape checking being associated with a 0.42 (p<.001) increase in concern about weight (see graph 3 in Figure 38). Whilst concern about shape decreased slowly throughout treatment in the control group, frequency of shape checking changed little (see Figure 36). In the intervention group, frequency of shape checking and concern about shape demonstrated the same trajectory, although there was no clear temporal precedence (see Figure 37). See Appendix 32 for Stata code and Appendix 35 for full model output.

7.4.3.4. Step 4: Impact of the Shape checking intervention on the Outcome (Concern about Shape) when Accounting for the Purported Mediator (Frequency of Shape Checking)

When including intervention status and frequency of shape checking (purported mediator) in the same model, intervention status was no longer significant (p=.795). The impact of the intervention on the outcome (concern about shape) decreased from -0.09 points each intervention week to 0.01 (see graph 4 in Figure 38). This represents a total reduction in the effect of intervention week on the outcome. There was a significant interaction between the shape checking intervention and the purported mediator (frequency of shape checking), suggesting that changes in frequency of checking had a greater effect on concern about shape in patients in the shape checking intervention group. This suggests that frequency of shape checking explained most of the variance in outcomes in the shape checking intervention group,

supporting the hypothesis that frequency of shape checking is a potential mediator of the effect of treatment on concern about shape in CBT-E. See Appendix 32 for Stata code and Appendix 36 for full model output.

Lagged models assessing the effect that prior level of the purported mediator had on subsequent outcome found little evidence of a relationship between earlier measures of frequency of shape checking (purported mediator) and the outcome (concern about shape).

7.4.4. Study 3b: An In-Depth Analysis using Autoregressive Structural Equation Modelling for Statistical Mediation

To verify the results of study 3a, namely that there was an indirect relationship between the shape checking intervention and concern about shape via frequency of shape checking in CBT-E, an autoregressive structural equation model was used to directly model the relationships between treatment, frequency of shape checking and concern about shape. As described in the analytic considerations, a shared time line was created to enable comparisons between the intervention group and the control group. Treatment allocation was dropped from the model due to the high percentage of overlap with intervention status. Results from study 3a corroborated this decision; treatment allocation did not significantly predict the outcome (concern about shape). Nonetheless, study 3b also suggested that the purported mediator (frequency of shape checking) increased in patients in CBT-E when controlling for the effect of shape checking intervention. This means that study 3b estimates for the effect of the shape checking intervention on the purported mediator are potentially conservative.

7.4.4.1. Effect of the Shape Checking Intervention on the Purported Mediator (Frequency of Shape Checking)

There was a significant difference between treatments, with frequency of shape checking (purported mediator) increasing in the intervention group by 0.40 points more than the control group in the first week of the intervention (see Figure 39). This was followed by significant decreases in frequency of shape checking of -0.49 and -0.45 points in intervention weeks 3 and 6 respectively, with an overall negative trend in all other weeks except week 5. However, the strength of relationship of prior frequency of shape checking to frequency of shape checking at later time points did not vary significantly during this period; X^2 (5, N = 130) = 8.93, *p*=.112, suggesting levels remained more or less stable during this period. This suggests that frequency of shape checking temporarily increased during the shape checking 'monitoring' stage of the intervention, and subsequently decreased in the shape checking 'addressing' stage of the intervention.

7.4.4.2. Effect of the Shape Checking Intervention on the Outcome (Concern about Shape) when Accounting for the Purported Mediator (Frequency of Shape Checking)

There were no significant differences in concern about shape (outcome) between the intervention group and the control group at any intervention week when accounting for the purported mediator (see Figure 39). However, the strength of relationship of prior concern about shape to concern about shape at later time points varied significantly during this period; X^2 (5, N = 130) = 14.61, *p*=.012, suggesting that levels of concern about shape fluctuated over time. This suggests that intervention status did

not have a significant effect on the outcome (concern about shape) when the purported mediator (frequency of shape checking) was accounted for.

7.4.4.3. Indirect Effect of the Shape Checking Intervention on the Outcome (Concern about Shape) via the Purported Mediator (Frequency of Shape Checking)

There was a moderate correlation between frequency of shape checking (purported mediator) and concern about shape (outcome) at the start of the intervention period (see Figure 39). Whilst these variables continued to be associated through this period, with the exception of intervention weeks 1 and 2, the effects of the previous week and the current week combined to cancel each other out. There was no evidence for the shape checking intervention reducing concern about shape indirectly via a decrease in frequency of shape checking (see Table 16). However, there was a significant indirect effect for increasing concern about shape via frequency of shape checking at intervention week 1 only when frequency of shape checking increased. This suggests that in the short-term frequency of shape checking was related to change in concern about shape, but not in the direction hypothesised. In the long-term, decreases in frequency of shape checking were unrelated to change in concern about weight i.e. the purported mediator did not mediate the effect of CBT-E on the outcome. See Appendix 37 for Stata code and Appendix 38 for full model output.

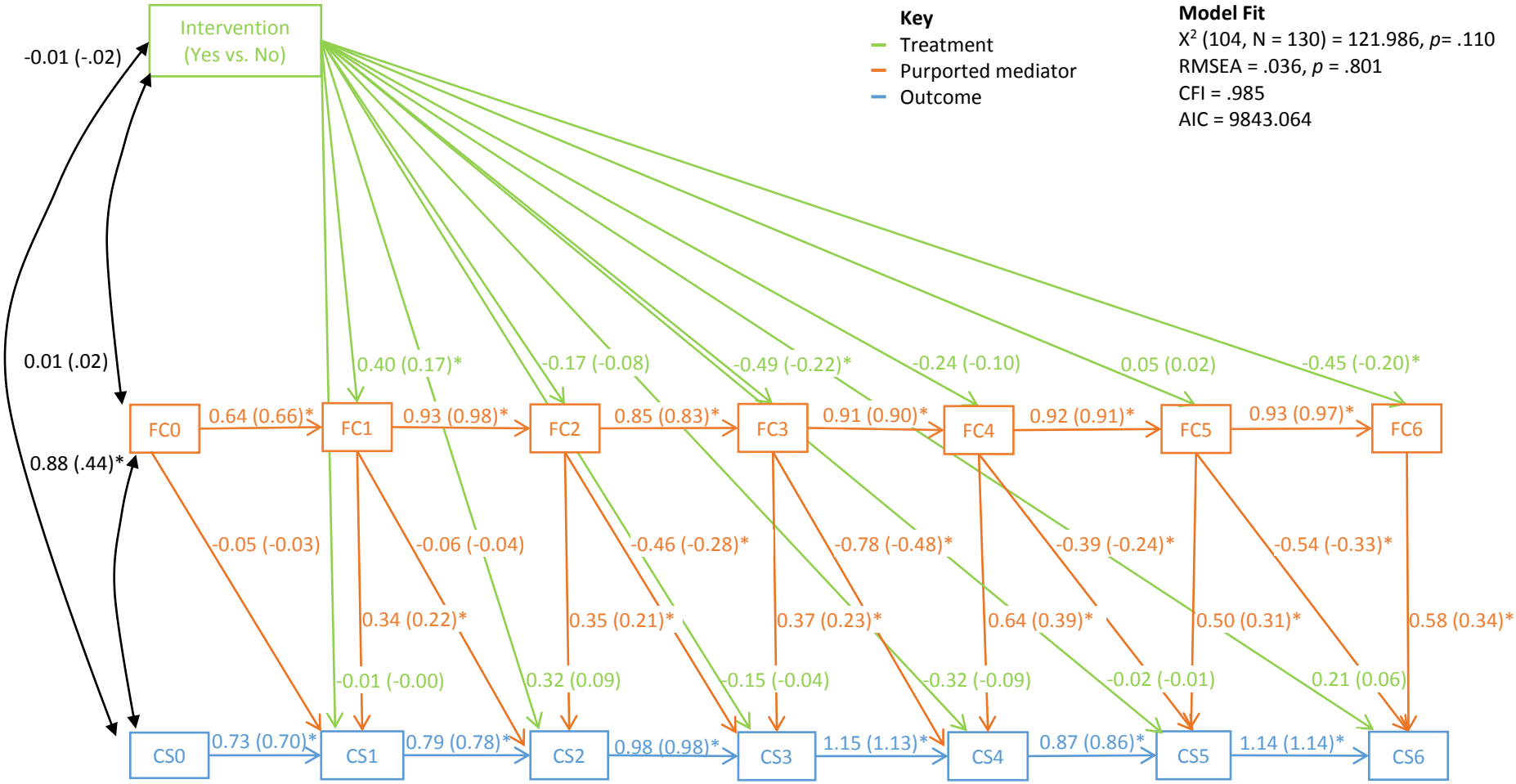


Figure 39: Autoregressive structural equation model of the impact of the shape checking intervention on concern about shape (CS) via frequency of shape checking, (FC). Numeral indicates intervention week. Controlling for shared measurement variance, treatment allocation, eating disorder duration, age, weight concern, self-esteem, depression, importance of shape, frequency of checking and worry about shape at baseline (standardised estimates in brackets). Significance denoted by * based on bootstrapped bias-corrected confidence intervals (1000 resamples). Missing data accounted for using full information maximum likelihood estimation.

Table 16: Comparison of the cumulative indirect and total effects of the shape checking intervention on concern about shape each week controlling for shared measurement variance, treatment allocation, eating disorder duration, self-esteem, age, weight concern, depression, importance of shape, frequency of checking and concern about shape at baseline

Impact of Shape Checking Intervention on Concern about Shape	Shape Checking Intervention Week	B	SE	Bias-Corrected 95% CI of B Coefficient		β
				Lower	**Upper**	
Indirect effects	1	0.14	0.08	0.03	0.36	0.04
	2	0.14	0.20	-0.27	0.50	0.04
	3	0.24	0.31	-0.31	0.95	0.07
	4	0.02	0.34	-0.61	0.68	0.01
	5	-0.28	0.28	-0.85	0.22	-0.07
	6	-0.61	0.40	-1.45	0.18	-0.15
Total effects	1	0.12	0.22	-0.34	0.53	0.04
	2	0.47	0.28	-0.06	1.08	0.13
	3	0.09	0.26	-0.41	0.64	0.03
	4	-0.31	0.29	-0.86	0.24	-0.07
	5	-0.31	0.34	-0.96	0.33	-0.08
	6	-0.40	0.36	-1.10	0.34	-0.10

Notes. B=unstandardized estimate, SE=standard error, CI= confidence interval, β=standardised estimate. CIs are bias-corrected based on 1000 bootstrap resamples. Indirect and total effects of treatment incorporate transmission of treatment effects via autoregressive paths (rather than each time point in isolation).

7.5. Summary of Results

The purported mediator (frequency of shape checking) decreased significantly in the intervention group in study 3a. However, study 3b indicated that this was preceded by a significant, temporary increase in frequency of shape checking during the monitoring stage of the shape checking intervention. Study 3a suggested that concern about shape (outcome) decreased in both treatments; however, patients in the intervention group experienced additional decreases in concern about shape following the implementation of the shape checking intervention. When accounting for the purported mediator (frequency of shape checking there was no longer a significant relationship between intervention status and the outcome (concern about shape). Whilst study 3a suggested a weak positive relationship between the purported mediator (frequency of shape checking) and the outcome (concern about shape), study 3b suggested that the effects of the previous week and the current week combined to cancel each other out at later weeks, providing little evidence for a consistent ongoing relationship between the two variables in the long-term. There was no evidence for the shape checking intervention decreasing concern about shape (outcome) via decreasing frequency of shape checking in study 3b. Instead, there was evidence for an indirect effect of the intervention on temporarily increasing concern about shape via increasing frequency of shape checking immediately following the implementation of the shape checking procedure.

Overall, the evidence for frequency of shape checking being a mediator of the effect of CBT-E on the outcome (concern about shape) was inconsistent.

Chapter 8. Study 4: Interpersonal Problem Severity as a Mediator of the Effect of Interpersonal Psychotherapy for Eating Disorders on Eating Disorder Psychopathology

8.1. Background

Rather than having multiple treatment components like enhanced cognitive behaviour therapy (CBT-E), interpersonal psychotherapy for eating disorders (IPT-ED) can be conceptualised as a complete unit that focuses predominantly on addressing interpersonal problems (Murphy et al., 2009). As part of this process the patient and therapist jointly identify and agree on one or two current interpersonal problems or "problem areas" (i.e. interpersonal role disputes, role transitions, grief, interpersonal deficits and life goals) to focus on during treatment that might be contributing to the maintenance of their eating disorder (see Chapter 1 for more detail).

8.2. Aim/Hypothesis

The interpersonal model of eating disorders proposes that interpersonal problems maintain the patient's eating disorder. It therefore hypothesises that helping patients to address their interpersonal problems will result in the resolution of their eating disorder psychopathology (see Figure 40).

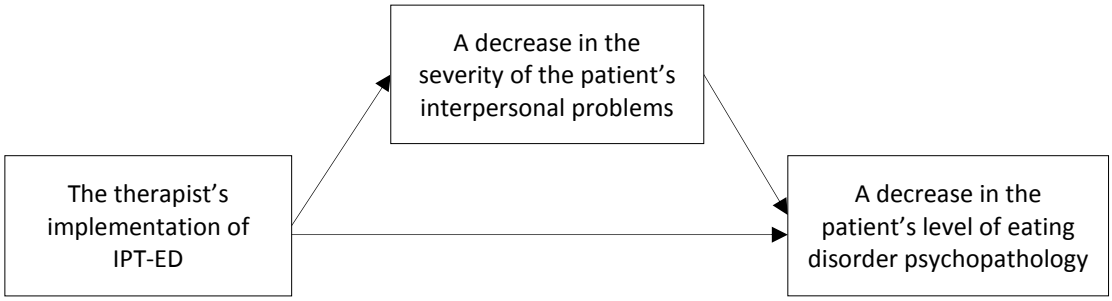


Figure 40: Illustration of the purported mediational relationship between interpersonal psychotherapy for eating disorders (IPT-ED), interpersonal problem severity and eating disorder psychopathology

In comparison, CBT-E does not explicitly target interpersonal problems. Although a reciprocal relationship might exist between eating disorder psychopathology and interpersonal functioning (whereby improvements in eating disorder psychopathology might improve interpersonal functioning), it would be hypothesised that IPT-ED would have a greater effect on interpersonal problems than CBT-E due to its greater focus on these problems in treatment. That being said, it takes considerable time for interpersonal difficulties to change as it often involves others changing too (Murphy et al., 2009). This may account for the differential time course of outcomes of CBT-E and IPT-ED. Consequently, it has been hypothesised that improvement in interpersonal problem severity may not occur until towards the end of treatment or even in the months afterwards (Murphy et al., 2009).

The aim of this study was to assess whether change in interpersonal problem severity mediates change in eating disorder psychopathology to a greater extent in IPT-ED compared to CBT-E. It was hypothesised that patients in IPT-ED would decrease their severity of interpersonal problems to a greater extent than patients in CBT-E and that this decrease would mediate change in eating disorder psychopathology.

8.3. Method

8.3.1. Experimental Procedure

Patients completed the 'Interpersonal Interview (IPI)' in treatment week 3, treatment week 19, and at 20-week and 40-week follow-up based on the 'Interpersonal Problem Form', which they completed with their therapist in week 3 prior to their initial interview. Instructions for completing the Interpersonal Problem Form differed depending on treatment arm. In IPT-ED, completion of the Interpersonal Problem Form was based on the interpersonal problems and interpersonal problem areas that the therapist and patient had agreed would be the focus of treatment. In CBT-E, its completion was based on a discussion with the patient about their current life and relationship difficulties. A period of three weeks into treatment was chosen to give therapists time to become better acquainted with their patients' and their life circumstances, and to match when in IPT-ED the decision is made regarding the interpersonal problem areas to be targeted in treatment. Interviews were undertaken by three independent assessors who were blind to treatment allocation. The interviewers underwent an initial training period in conducting and scoring the interpersonal interview before undertaking interviews themselves, and then attended monthly supervision together to ensure interviews continued to be conducted and scored in a consistent manner. Patients were warned not to disclose any information regarding their treatment to the interviewer to maintain treatment blinding. By consent, access to information from the 'Eating Disorder Examination (EDE) 16.0' was obtained from the main efficacy trial, which was collected at baseline, end of treatment, at 20-week follow-up and 40-week follow-up.

8.3.2. Study Measures

8.3.2.1. Interpersonal Problem Severity

Interpersonal problem severity was assessed using question 1 of the IPI (see Appendix 8). This question assessed the frequency and duration of the interpersonal problem identified in the Interpersonal Problem Form (see Appendix 9), and how much it had affected the patient over the 4 weeks prior to interview (or in the case of the interview at week 3 , the 4 weeks prior to starting treatment). Interpersonal problem severity was scored on a scale of 0 – 6, where 0 indicated no problem or the problem having been resolved, and 6 indicated a marked problem that was significantly interfering with the patient's life.

An inter-rater reliability study was undertaken as part of checking measure reliability across the three interviewers. A weighted Kappa was implemented using the following weightings; an exact match was coded as 1 (indicating full agreement), one score either side was coded as .80 (indicating 80% agreement) and anything else was coded as 0 (indicating no agreement)). Twelve interpersonal problems rated by the primary interviewer (EF) and the second interviewer (MOC), and six with the primary interviewer and the third interviewer (AD) were assessed. Analyses indicated substantial agreement between raters; Kappa = .86, bias-corrected 95% CI [.79, .93] and Kappa = .89, bias-corrected 95% CI [.85, .93], respectively (Landis & Koch, 1977).

8.3.2.2. Eating Disorder Psychopathology

Eating disorder psychopathology was assessed using the EDE. This interview gives an overall score assessing the severity of eating disorder psychopathology (dietary

restraint, eating, shape and weight concern) on a scale of 0 – 6, where 0 indicates no eating disorder psychopathology and 6 indicates extreme eating disorder psychopathology.

8.3.3. Analytic considerations

8.3.3.1. Potential Confounders

A potential unmeasured confounder was depression due to its hypothesised relationship with both interpersonal problems and eating disorder psychopathology. Depression was also potentially a mediator of the effect of treatment on the outcome given that interpersonal psychotherapy is an effective treatment of depression, and therefore would have been excluded from the analysis. However, as discussed in study 1, although depression was not measured, the influence of clinical depression on these variables were controlled for as those patients requiring treatment for depression were treated before entering the trial, and anti-depressant medication was maintained during treatment. Those who became depressed during treatment were put on hold while the depression was treated.

8.3.3.2. Exclusion of Confounded Non-Specific Predictors

Concern about weight and importance of shape baseline covariates were excluded as these items form part of the EDE.

8.4. Results

8.4.1. Non-Normality and Missing Data

Examination of the data suggested non-normality, with both interpersonal problem severity and eating disorder psychopathology being negatively skewed suggesting high levels of interpersonal problem severity and eating disorder psychopathology. There was a higher percentage of missing data on the interpersonal interview in CBT-E compared to IPT-ED (24.6% vs. 17.0%, respectively). There was no clear reason for this disparity. Missing data was mostly due to attrition.

8.4.2. Study 4a: A Preliminary Investigation using Multilevel Modelling for Statistical Mediation

8.4.2.1. Step 1: Impact of the Intervention (IPT-ED) on the Purported Mediator (Interpersonal Problem Severity)

Interpersonal problem severity (purported mediator) decreased in both IPT-ED and CBT-E (see Figure 41 and Figure 42). The interaction between treatment allocation and time was not significant (B=-0.01, p=.933), suggesting similar rates of change in the two treatments. Interpersonal problem severity decreased by -2.02 points (p<.001) every 20 weeks from the start of treatment (see graph 1 in Figure 43). This was qualified by a significant quadratic trend of 0.40 (p<.001) that suggested that change plateaued during follow-up, with little further change occurring after 20-week follow-up. The majority of change therefore occurred during treatment. See Appendix 39 for Stata code and Appendix 40 for full model output.

8.4.2.2. Step 2: Impact of the Intervention (IPT-ED) on the Outcome (Eating Disorder Psychopathology)

Eating disorder psychopathology (outcome), as measured by the EDE, decreased in both IPT-ED and CBT-E (see Figure 41 and Figure 42). Overall, eating disorder psychopathology decreased at a rate of -1.92 points every 20 weeks (see graph 2 in Figure 43). This was qualified by a significant quadratic trend of 0.44, suggesting that rate of change slowed towards the end of follow-up. There was a significant interaction between treatment allocation and time. In IPT-ED, eating disorder psychopathology was 0.59 points (p=.034) higher than CBT-E at each time point. This was qualified by a significant quadratic trend of -0.16 (p=.060, bias-corrected CI [-0.34, -0.01]), suggesting that the difference between treatments decreased over follow-up. See Appendix 39 for Stata code and Appendix 41 for full model output.

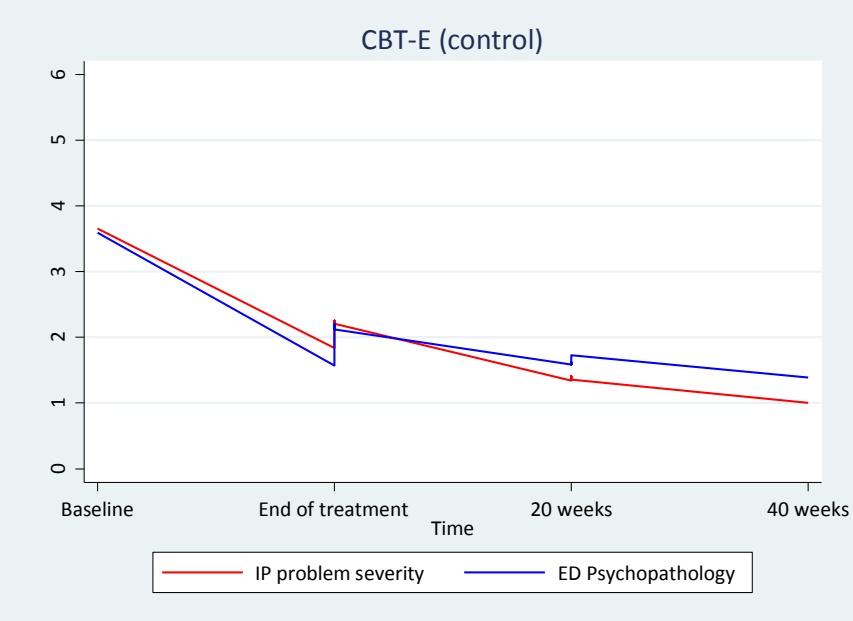


Figure 41: Line graph plotting the observed means (smoothed curve) of interpersonal problem severity and eating disorder psychopathology in CBT-E during treatment and follow-up

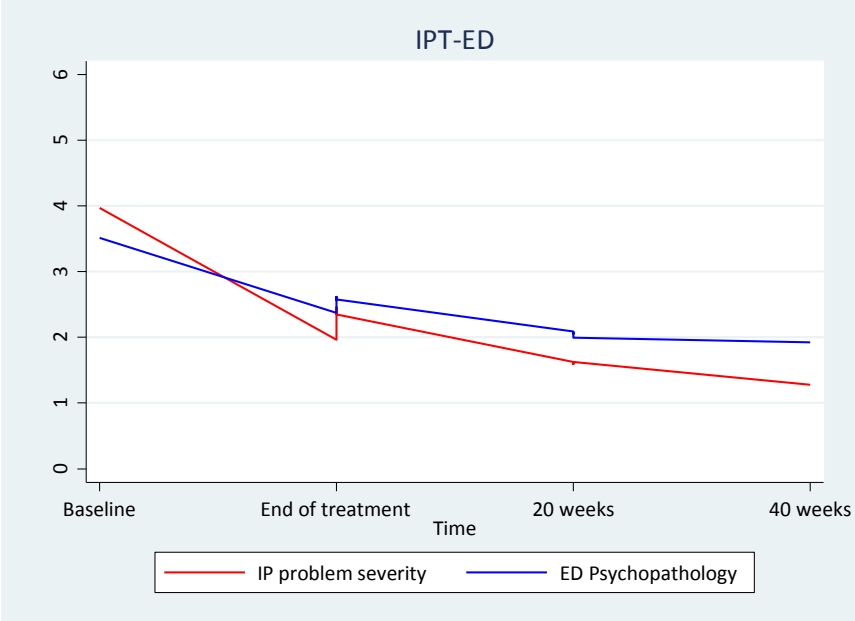


Figure 42: Line graph plotting the observed means (smoothed curve) of interpersonal problem severity and eating disorder psychopathology in IPT-ED during treatment and follow-up

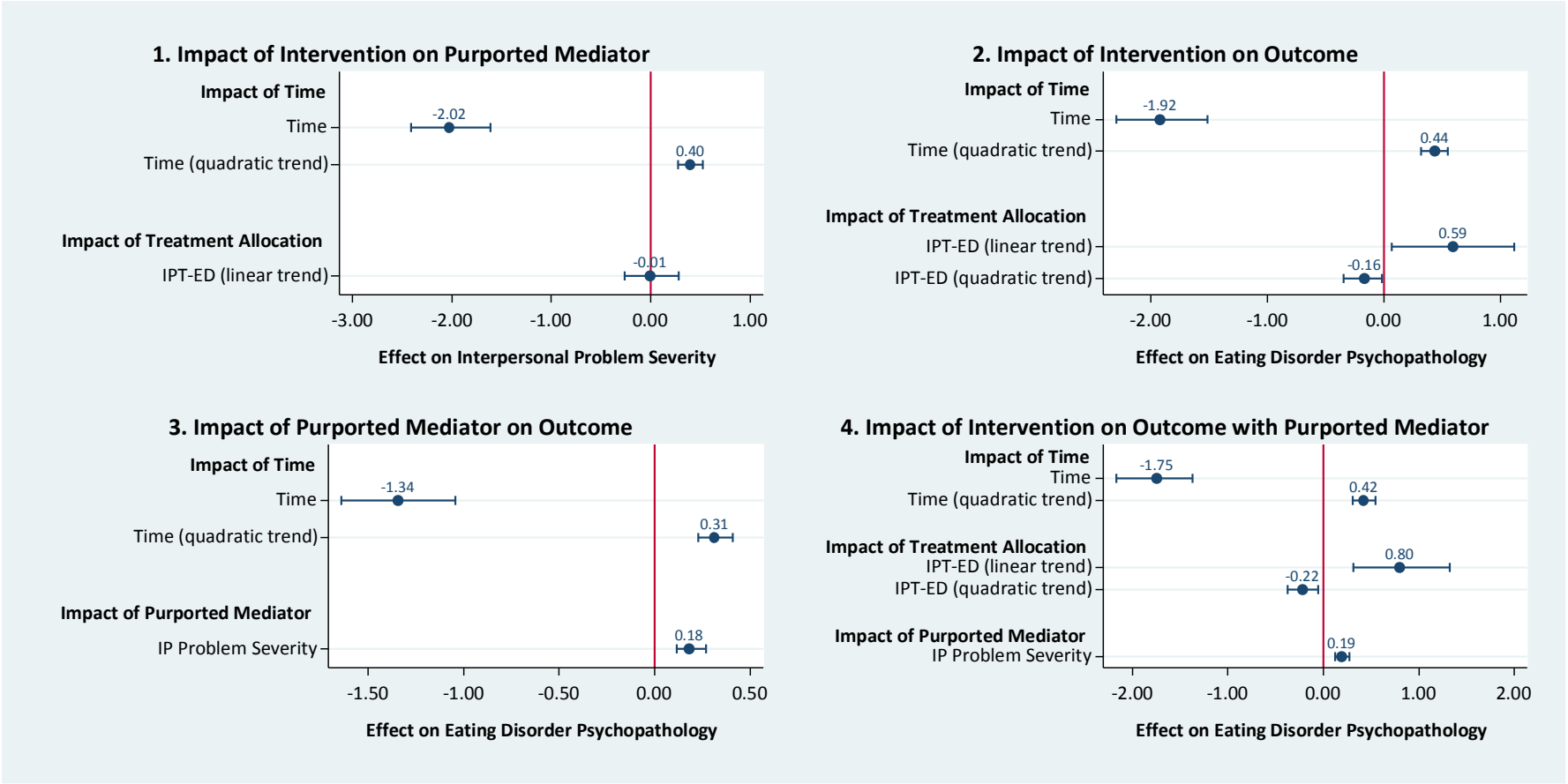


Figure 43: Coefficient plots with 95% bias-corrected confidence intervals (1000 bootstrap resamples) comparing the effect of IPT-ED on interpersonal problem severity and eating disorder psychopathology when interpersonal problem (IP) severity was / was not included as a mediator. Impact of treatment allocation represents an interaction between treatment allocation and treatment week. Controlling for the effects of eating disorder diagnosis, eating disorder duration, self-esteem and depression at baseline.

8.4.2.3. Step 3: Relationship between the Purported Mediator (Interpersonal Problem Severity) and the Outcome (Eating Disorder Psychopathology)

Interpersonal problem severity (purported mediator) was positively associated with eating disorder psychopathology, with a one point increase in interpersonal problem severity being associated with a 0.18 (p<.001) increase in eating disorder psychopathology (see graph 3 in Figure 43). There were few differences in the trajectories of these two variables between the two treatments, although, in IPT-ED change in interpersonal problem severity appeared greater than that in eating disorder psychopathology, and was vice versa in CBT-E (see Figure 41 and Figure 42). However, there was no clear temporal precedence of change in either of the two treatments. See Appendix 39 for Stata code and Appendix 42 for full model output.

8.4.2.4. Step 4: Impact of IPT-ED on the Outcome (Eating Disorder Psychopathology) when Accounting for the Purported Mediator (Interpersonal Problem Severity)

When including treatment allocation and interpersonal problem severity (purported mediator) in the same model, the impact of treatment allocation on the outcome (eating disorder psychopathology) remained significant (p= .002), and the coefficient increased by 0.21 points. This represents a 36% increase in effect (see graph 4, Figure 43). This suggests that interpersonal problem severity explained some of the variance in outcome in the two treatments, and that interpersonal problem severity might contribute to the differential outcome of the two treatments, supporting the hypothesis that interpersonal problem severity might be a potential mediator of the

effect of IPT-ED on the outcome. See Appendix 39 for Stata code and Appendix 43 for full model output.

Lagged models assessing the effect that prior level of the purported mediator had on subsequent outcome found little evidence of a relationship between earlier measures of interpersonal problem severity (purported mediator) and the outcome (eating disorder psychopathology).

8.4.3. Study 4b: An In-Depth Analysis using Autoregressive Structural Equation Modelling for Statistical Mediation

To verify the results of study 4a, namely that there was no indirect relationship between treatment and eating disorder psychopathology via interpersonal problem severity in IPT-ED, an autoregressive structural equation model was used to directly model the relationships between treatment, interpersonal problem severity and eating disorder psychopathology.

8.4.3.1. Effect of IPT-ED on the Purported Mediator (Interpersonal Problem Severity)

There were no significant differences between the two treatments in terms of interpersonal problem severity at any time point (see Figure 44). Furthermore, the strength of relationship of prior interpersonal problem severity to interpersonal problem severity at later time points did not vary significantly during this period, suggesting that interpersonal problem severity remained stable during this period; X^2 (2, N = 130) = 0.35, p = .839. This suggests that both IPT-ED and CBT-E significantly decreased the purported mediator (interpersonal problem severity) during this period.

8.4.3.2. Effect of IPT-ED on the Outcome (Eating Disorder Psychopathology) when Accounting for the Purported Mediator (Interpersonal Problem Severity)

There was a significant difference between the treatments in eating disorder psychopathology (outcome), with patients in IPT-ED scoring 0.83 points higher on eating disorder psychopathology than patients in CBT-E (see Figure 44). There were no significant differences at later time points, however. The strength of relationship of prior eating disorder psychopathology to eating disorder psychopathology at later time points remained more or less stable during this period; X^2 (2, N = 130) = 1.32, p = .517. This suggests that treatment allocation had a significant effect on the outcome (eating disorder psychopathology), with IPT-ED performing worse than CBT-E, even when accounting for the purported mediator (interpersonal problem severity).

8.4.3.3. Indirect Effect of IPT-ED on the Outcome (Eating Disorder Psychopathology) via the Purported Mediator (Interpersonal Problem Severity)

Overall there was weak positive relationship between interpersonal problem severity and eating disorder psychopathology, except at 20 week follow-up when the combined effects of previous and current interpersonal problem severity cancelled each other out (this is when interpersonal problem severity would have begun to stabilise) (see Figure 44). Although there was evidence of an indirect effect of IPT-ED having increased eating disorder psychopathology at 20-week follow-up (see Table 17) this was likely via the direct effect of treatment allocation on eating disorder psychopathology at end of treatment, rather than its effect via interpersonal problem

severity. The evidence for IPT-ED reducing eating disorder psychopathology via a decrease in interpersonal problem severity was inconclusive. See Appendix 44 for Stata code and Appendix 45 for full model output.

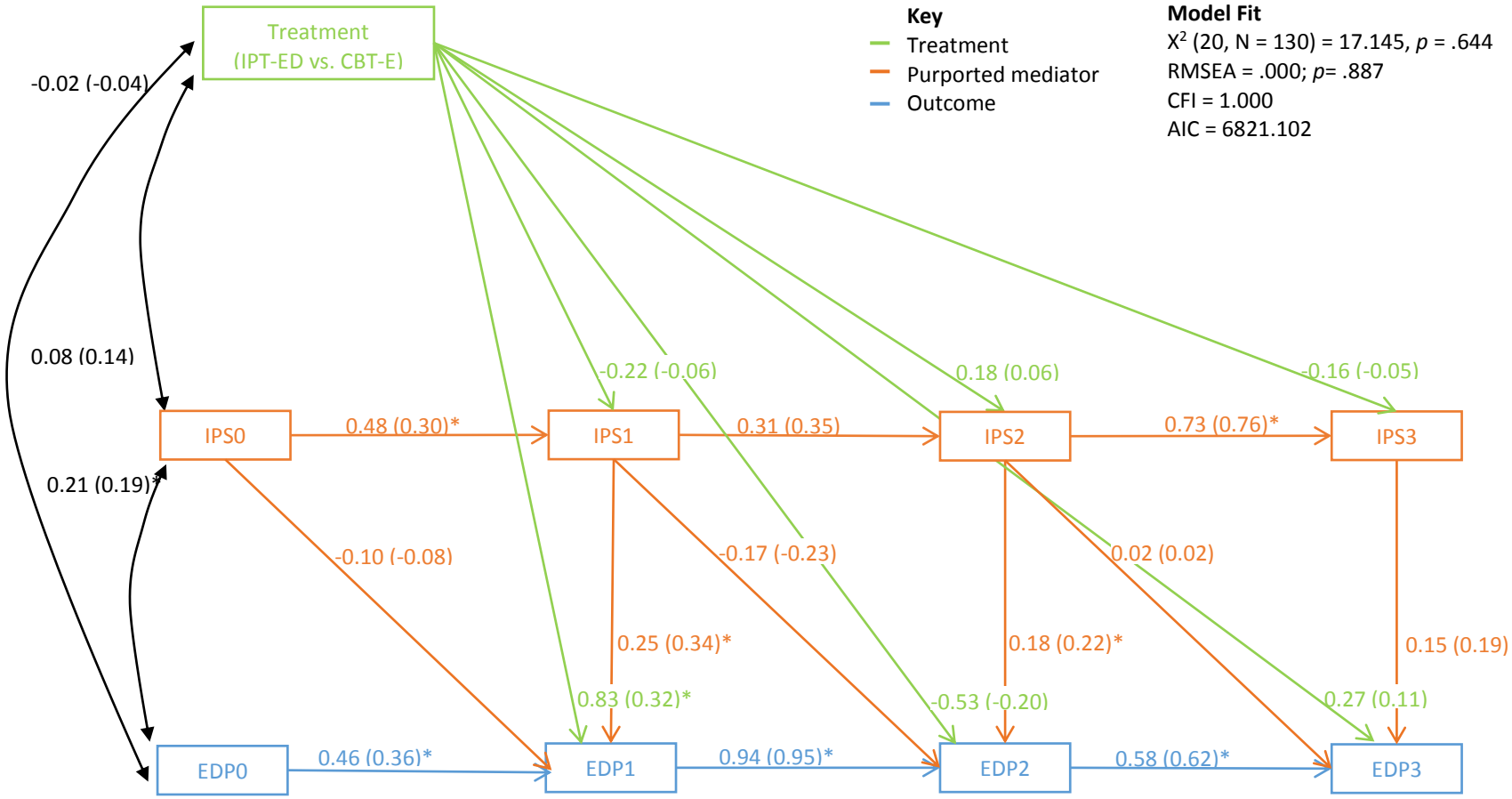


Figure 44: Autoregressive structural equation model of the impact of treatment on eating disorder psychopathology (EDP) via interpersonal problem severity (IPS). Numeral indicates study time point (0=baseline, 1=end of treatment, 2=20 week follow-up, 3=40 week follow-up). Controlling for shared measurement variance, eating disorder diagnosis, eating disorder duration, self-esteem, and depression at baseline (standardised estimates in brackets). Significance denoted by * based on bootstrapped bias-corrected confidence intervals (1000 resamples). Missing data accounted for using full information maximum likelihood estimation.

Table 17: Comparison of the cumulative indirect and total effects of IPT-ED on interpersonal problem severity at each study time point controlling for shared measurement variance, eating disorder duration, depression and self-esteem at baseline

Impact of Treatment Allocation on Eating Disorder Psychopathology	Study Time Point	B	SE	Bias-Corrected 95% CI of B Coefficient		β
				Lower	**Upper**	
Indirect effects	End of treatment	-0.05	0.10	-0.30	0.12	-0.02
	20-week follow-up	0.79	0.35	0.31	1.80	0.30
	40-week follow-up	0.13	0.16	-0.15	0.50	0.06
Total effects	End of treatment	0.78	0.23	0.31	1.24	0.30
	20-week follow-up	0.25	0.23	-0.15	0.72	0.10
	40-week follow-up	0.41	0.22	-0.04	0.83	0.17

Notes. B=unstandardized estimate, SE=standard error, CI= confidence interval, β=standardised estimate. CIs are bias-corrected based on 1000 bootstrap resamples. Indirect and total effects of treatment incorporate transmission of treatment effects via autoregressive paths (rather than each time point in isolation).

8.5. Summary of Results

The purported mediator (interpersonal problem severity) decreased significantly in both CBT-E and IPT-ED at a similar rate. This was consistent in both studies. Whilst eating disorder psychopathology (outcome) decreased significantly in both treatments in study 4a, it decreased more slowly in IPT-ED compared to CBT-E. When accounting for the purported mediator (interpersonal problem severity), which was significantly positively related to the outcome (eating disorder psychopathology), the relationship between treatment allocation and poorer outcome increased in strength. Study 4b supported a significant difference between treatments, with IPT-ED having poorer outcomes at the end of treatment. Whilst there was a significant relationship between interpersonal problem severity and eating disorder psychopathology, there was no evidence for IPT-ED decreasing eating disorder psychopathology (outcome) via decreasing interpersonal problem severity (purported mediator).

Overall, evidence for interpersonal problem severity as a potential mediator of the effect of IPT-ED on eating disorder psychopathology was inconclusive.

Chapter 9. Discussion

9.1. Overview

The aim of this thesis was to investigate whether the hypothesised mechanisms of action of enhanced cognitive behaviour therapy (CBT-E) and interpersonal psychotherapy for eating disorders (IPT-ED) account for the effect that treatment has on patients' outcomes. Whilst cognitive behaviour therapy (CBT) is currently the leading treatment for non-underweight eating disorders, and interpersonal psychotherapy (IPT) is a leading competitor, up to a third of patients in CBT-E and half of patients in IPT-ED do not improve following treatment. There is therefore a need to increase their potency. Such work has been impeded by a lack of knowledge regarding how these treatments operate and whether they work as hypothesised. Accordingly, the candidate analysed data from a randomised controlled trial (RCT) comparing CBT-E and IPT-ED, which had a series of mediation studies embedded within it. This trial is the first efficacy and mechanism evaluation trial of CBT-E compared to IPT-ED, and the first ever to examine purported mediator/s of IPT. Three of the key treatment procedures in CBT-E were examined; "regular eating", "weekly weighing" and "shape checking", and what is arguably the key procedure in IPT-ED; addressing interpersonal problems. An analytic strategy was developed combining several existing statistical methods in a novel way to assess for statistical mediation when comparing two active treatments. An approach for incorporating treatment personalisation into statistical mediation analyses was also developed, an issue which has received little prior

attention. The findings therefore inform both clinical understanding of these treatments, and provide guidance for undertaking statistical mediation when two active treatments are being compared, and for incorporating treatment personalisation into statistical mediation analyses. The clinical and statistical implications of this thesis will be discussed separately.

9.2. Key Findings and Implications

9.2.1. Purported Mediators of Enhanced Cognitive Behaviour Therapy

9.2.1.1. Study 1: Regularity of Eating as a Mediator of the Effect of CBT-E on Frequency of Binge Eating (Outcome)

The results are consistent with the hypothesis that regularity of eating is a mediator of the effect of CBT-E on frequency of binge eating. However, as temporal precedence of the purported mediator could not be established, the results are not definitive. It could be argued that change in the purported mediator (regularity of eating) was the result of change in the outcome. This is not supported by these findings as modelling frequency of binge eating (the outcome) as an alternative mediator did not affect the relationship between treatment and regularity of eating (the purported mediator). This finding supports the direction of influence flowing from regularity of eating to frequency of binge eating. An alternative explanation for these findings could be increased quantity of food consumed between binges, particularly carbohydrates. However, this is unlikely as patients were instructed to change only the timing of their eating, not what they were eating or how much they ate (see Fairburn, Cooper,

Shafran, et al., 2008b). The results are consistent with studies examining cognitive behavioural treatments for eating disorders, which found similar relationships between dietary restraint and frequency of binge eating (Ellison et al., 2016; Shah et al., 2005; Waller et al., 2013; Wilson et al., 2002; Zendegui et al., 2014). They are also consistent with research examining the relationship between dietary restraint and frequency of vomiting (as a proxy for binge eating) (Spangler et al., 2004). With regard to the timing of change, the current study replicates previous findings from an RCT comparing cognitive behaviour therapy for bulimia nervosa (CBT-BN) and interpersonal psychotherapy for bulimia nervosa (IPT-BN). This study compared change in frequency of binge eating in these two treatments although regularity of eating was not examined (R. Jones, Peveler, Hope, & Fairburn, 1993). A rapid decrease in the frequency of binge eating was observed in CBT-BN, with the change in IPT-BN being slightly delayed in comparison.

Although regularity of eating did appear to mediate a reduction in frequency of binge eating (outcome) in CBT-E, the decrease in frequency of binge eating was modest (approximately 1.40 binge eating episodes a week during the first four weeks of treatment). Given that regularity of eating changed little in IPT-ED yet frequency of binge eating decreased significantly by the end of the first four weeks of treatment. This suggests that change in frequency of binge eating was mediated via a different process in IPT-ED. This fits with the different theoretical conceptualisations of these two treatments. Whilst, it is not possible to know whether regularity of eating was maintained in later weeks of treatment, or continued to be associated with frequency of binge eating, other research supports the longer term effects of the regular eating

procedure. In an RCT comparing CBT-BN and IPT-BN, change in regularity of eating at week four of treatment was related to a lower frequency of binge eating at the end of treatment (Wilson et al., 2002). Presence of binge eating in the month prior to starting treatment was tested as an additional moderator as it was possible that the regular eating procedure might have differential effects in patients who did / did not binge eat. This did not appear to be the case in this study.

The results suggest that the majority (80%) of change in regularity of eating (purported mediator) occurred within the first week of treatment in CBT-E. This suggests that change in regularity of eating occurred much faster than originally anticipated. Given that 'early treatment response' is predictive of later outcome it therefore might be important to pay attention to and target patients who do not experience improvement during this period. A better understanding of the time course might help identify patients who are struggling with treatment.

The fact that temporal precedence could not be established could be related to non-optimal timing of the measurements. Non-optimal timing of measurements might explain the contradictory relationships found between regularity of eating and frequency of binge eating identified in study 1b. Contemporaneous paths between these variables (i.e. relationships within the same treatment week) indicated a negative relationship as hypothesised. However, lagged paths between these two variables (relationships from one treatment week to the next) suggested a positive relationship, although overall a negative relationship was found. It is possible that this may be indicative of lapses in regularity of eating / binge eating between each treatment week but that the measurement intervals are too wide to detect these

fluctuations. At this point in treatment sessions occur on bi-weekly basis, therefore patients are receiving 'two doses' of therapy in the space of one measurement period. It could also be that the relationships between change in regularity of eating and change in frequency of binge eating change occurs more rapidly than can be observed on a session by session basis

Ecological momentary assessment research, in which patients' thoughts, feelings and behaviours are randomly assessed throughout the day over a period of days and weeks, supports this. One study in women with bulimia nervosa suggests that a dynamic relationship exists between dieting (e.g. caloric restriction) and frequency of binge eating, with dietary restriction the day before impacting on binge eating the subsequent day (Zunker et al., 2011). A second study in women with typical and atypical anorexia nervosa suggested that there is a cumulative effect over time in those who had experienced an episode of binge eating the day before compared to those who had not, with those who fasted over several consecutive days binge eating more frequently (De Young et al., 2014). This suggests that there may be both dynamic and cumulative processes involved in the relationship between regularity of eating and frequency of binge eating, which could be better examined using daily self-monitoring records or ecological momentary assessment.

Future research should examine change in regularity of eating and frequency of binge eating within the first two weeks of this intervention to see if temporal precedence can be established when examining a shorter time period. Using session-by-session or daily self-monitoring, or ecological momentary assessment could help with this. It would also be useful to address the question as to what extent the CBT procedure of

regular eating needs to be implemented in order to affect frequency of binge eating, should this be a mediator. A dismantling design in which the regular eating procedure was included / excluded could also help assess the impact of this procedure on the outcome. However, such a design would require omitting a procedure that has been shown to be associated with outcome during and at the end of treatment, and therefore might have ethical implications. Future research should also consider investigating whether this early change in frequency of binge eating in IPT-ED is related to early changes in interpersonal functioning in IPT-ED following the identification of interpersonal problem areas. A similar assessment of interpersonal functioning in CBT-E would also help to give a broader understanding of these potentially differing processes in these two treatments.

9.2.1.2. Study 2: Frequency of Weighing as a Mediator of the Effect of CBT-E on Concern about Weight (Outcome)

The results do not support frequency of weighing as a mediator of change in concern about weight in CBT-E. Whilst frequency of weighing and concern about weight were associated at baseline, there was little evidence of an ongoing association between these variables once the weekly weighing procedure was implemented. It could be that increased frequency of weighing is more a manifestation of concern about weight than a maintaining factor of it.

These results do not necessarily suggest, that weekly weighing does not have a role to play in CBT-E. It has been argued that weekly weighing serves several purposes; ensuring patient safety by monitoring changes in weight, a way of identifying changes in eating pattern, a way of addressing anxiety associated with weighing and knowing

one's weight, and addressing the patient's erroneous beliefs regarding the direct link between eating and weight gain (Waller & Mountford, 2015). Reducing frequency of weighing may address other barriers that may impede successful treatment (e.g. anxiety, low self-esteem) that were not measured in this study. Furthermore, whilst a decreased frequency of weighing (purported mediator) may not relate to a decrease in concern about weight, it may still serve a role in preventing concern about weight from increasing following changes to eating e.g. regularity of eating, incorporating avoided foods into one's diet. It may also play a role in reinforcing decisions to diet and the patients' use of other methods of weight control.

The weekly weighing procedure is a complex intervention with several different aspects, including; decreasing frequency of weighing by prescribing abstinence from weighing between sessions, tracking weight and improving interpretation of weight and the causes of fluctuations in body weight. This study measured only the behavioural aspect of this intervention (frequency of weighing), which was used as a proxy for the impact of the whole intervention. It therefore did not measure the cognitive aspect of the intervention (how the patient interprets weight on the scale). These cognitive aspects might be more important in decreasing worry about weight. Changes in cognitive interpretation of weight could explain why the association between frequency of weighing and concern about weight appeared to decrease during the course of treatment. Even if patients are weighing themselves outside treatment sessions but are interpreting their weight correctly then small changes in weight would not necessarily increase their concern about weight. It is also possible that more dynamic processes were at play (as demonstrated by the positive

relationship observed within the same week, and the negative longitudinal relationship between consecutive weeks between frequency of weighing and concern about weight).

Frequent weigher status was assessed as an additional moderator in this study. It was hypothesised that the weekly weighing intervention might have a greater impact on patients who weigh themselves more frequently, and therefore might experience greater shifts in concern about weight fuelled by their focus on natural daily weight fluctuations. There was no evidence for this despite frequent weigher status being a moderator of the effect of treatment on frequency of weighing. This further supports the idea that frequent weighing was not related to greater concern about weight. Nonetheless, there may have been important differences between patients who were actively avoiding weighing themselves at the time treatment started and the rest of the sample. In this group of patients implementation of the weekly weighing procedure would have resulted in an increase in weighing frequency (an increase from zero to once per week in treatment). This may have resulted in a temporary increase in worry due to the patients' being exposed to something they had been avoiding. Unfortunately, numbers were too few to test this group separately from the main sample, but such a group could be the focus of further study in the future.

Future research should consider other measurements of the purported mediator, including assessments tapping into the cognitive processes theorised to underpin this intervention. A more in depth of analysis of weighing and worry about weight looking at immediate experiences around weighing using ecological momentary assessment methods asking a) fear about being weighed b) belief in accuracy of weight on scale c)

concern about weight d) number of times weighed that week (outside of the treatment session) could help to elucidate this. It might also help to identify key processes and help gain further insight into the relationship between frequency of weighing, cognitive processes around weighing and concern about weight. A qualitative exploration of patients' experiences of this procedure might also increase understanding of the impact of the weekly weighing procedure on concern about weight.

9.2.1.3. Study 3: Frequency of Shape Checking as a Mediator of the Effect of CBT-E on Concern about Shape (Outcome)

Overall, the evidence for frequency of shape checking being a mediator of the effect of CBT-E on concern about shape was inconsistent. There was limited evidence for CBT-E decreasing concern about shape (outcome) via decreasing frequency of shape checking; instead there was evidence for CBT-E temporarily increasing concern about shape via increasing frequency of shape checking immediately following the implementation of the shape checking procedure.

Temporary increases in concern about shape have been hypothesised as a potential side effect of monitoring shape checking behaviour during the shape checking procedure (Fairburn, 2008). It has been reported that many patients are unaware of the frequency that they check their shape, and bringing these behaviours into conscious awareness can be distressing (Fairburn, 2008). The temporary increase in concern about shape is consistent with this. Harder to explain is the finding that frequency of shape checking increased following the implementation of the shape checking procedure. However, it is possible that the reported increases in frequency

of shape checking may be due more to an increased awareness of these behaviours, as opposed to an actual increase in them. If that is the case, then the current analyses may not truly reflect the impact of addressing frequent shape checking in CBT-E.

Although frequency of shape checking and concern about shape were associated at baseline, there was little evidence for a consistent ongoing association between these variables once the shape checking procedure had been implemented. This might suggest that frequent shape checking is more a manifestation of concern about shape, rather than a maintaining factor of it. However, this notion is inconsistent with the finding that shape checking monitoring preceded an increase in concern about shape, suggesting that shape checking does affect concern about shape. There may have been more dynamic processes at play (as demonstrated by the positive relationship observed within the same week, and the negative longitudinal relationship between consecutive weeks) that were not captured by weekly measurement intervals.

Frequency of shape checking was not tested as a potential moderator in this analysis. One reason for this is that there was no obvious cut-off to use for defining frequent shape checkers, unlike in study 2 where frequent weighers were defined based on the frequency of weighing undertaken in CBT-E (i.e. once per week). Whilst this could have resulted in floor effects in the analysis, this is unlikely as daily shape checking is a common behaviour in both the general population and in patients with eating disorders (Kachani, Brasiliano, Cordás, & Hochgraf, 2013; Mintz et al., 2013; Shafran, Fairburn, Robinson, & Lask, 2004). Furthermore, frequency of shape checking was rated on a Likert scale assessing frequency over the last 7 days. This scale ranged higher than is assessed in the Eating Disorder Examination ("many times a day" vs.

"every day"). Very few patients in the sample did not check their shape (N=9, 7%) and the majority of patients checked their shape at least once a day (N=82, 63%). Frequent shape checking was therefore common in most patients in the sample.

The structural equation model focused only on the first six weeks of the shape checking procedure. However, it has been hypothesised that over-evaluation of shape and weight takes time to change (Fairburn, 2008). Therefore 6 weeks only may be too short a time period to detect change in concern about shape related to decreases in shape checking, particularly, given the initial temporary increases in shape concern following implementation of shape checking monitoring. The effect of decreasing frequency of shape checking may be cumulative, occurring over a longer period than was able to be examined in the structural equation model without it becoming overly complex.

This study measured both frequency of mirror checking and frequency of body checking (touching / measuring). Whilst these two items exemplify body scrutiny, and are strongly associated, it is possible that they capture slightly different behaviours, as mirror checking might be considered a more normative behaviour than body checking (particularly measuring). Furthermore, they failed to capture the cognitive aspects related to this intervention (e.g. normalisation of shape checking, bias towards disliked body parts).

Comparison making (comparing one's body to others') is also a form of shape checking that is addressed in CBT-E, but it was not included in this analysis for several reasons. Firstly, comparison making is seen as a special form of shape checking in that it occurs in a social context, rather than consisting of direct self-inspection of one's body.

Secondly, comparison making is addressed by getting patients to a) compare shape-neutral features, b) scrutinise other people's bodies to understand the effect of scrutiny on perception, or c) compare to a broader selection of people, rather than a biased sample of thin, attractive individuals. This means that patients may not necessarily report a decrease in comparison making, as it is the quality, not the quantity of comparison making that is changing, meaning that frequency of comparison making would not be appropriate as a measure.

This study did not consider the role of body avoidance in the analysis. However, body avoidance might be considered as the other end of the spectrum from frequent shape checking (Fairburn, 2008). Many patients report alternating between periods of frequent checking and avoidance (Shafran et al., 2004). It is possible that these processes may also be interrelated during treatment, particularly, in an intervention that attempts to decrease frequent checking. Therefore, body avoidance may represent an unmeasured confounder in this study.

Further consideration should also be given to the unique versus combined contributions of the different procedures used in the CBT-E "shape concern" module (of which frequency of body checking is only one). This module also includes increasing marginalised domains, addressing other forms of shape checking such as comparison making, body avoidance, feeling fat, and understanding the origins of concern about shape and the eating disorder mindset. It is possible that these procedures may also explain the decrease in concern about shape seen in this study.

Overall, there is not sufficient evidence to support refining or eliminating the shape checking procedure from CBT-E; instead, further research is required to better

understand this procedure. Research should examine how shape checking relates to shape controlling behaviours. It should also assess how frequency of shape checking and body avoidance relate during treatment, and control for this in analyses. Comparison making should be investigated as a separate form of shape checking. The cognitive aspects of the intervention and how these might relate to concern about shape should also be investigated. Subsequent studies should also consider how these variables are measured, both in terms of timing of measurement and what is being measured. One option would be to include multiple assessments using active monitoring (in both treatments) to overcome potential issues with heightened awareness artificially increasing self-reported frequency of checking and concern about weight. This would need to be done carefully to avoid overly distressing patients. Ecological momentary assessment might also be a good alternative to avoid prolonged active monitoring of shape checking behaviours by the patient, and might also capture more dynamic relationships between these variables.

9.2.2. Purported Mediators of Interpersonal Psychotherapy for Eating Disorders

9.2.2.1. Study 4: Interpersonal Problem Severity as a Mediator of the Effect of IPT-ED on Eating Disorder Psychopathology (Outcome)

Overall, the results were inconclusive regarding whether interpersonal problem severity was a potential mediator of the effect of IPT-ED on eating disorder psychopathology. As both interpersonal problem severity and eating disorder psychopathology had changed significantly before the end of treatment, temporal precedence of change could not be established. It is therefore possible that change in

interpersonal problem severity was a mediator in eating disorder psychopathology in either, or both, treatments.

The fact that temporal ordering could not be established suggests that change in interpersonal problem severity occurs earlier in IPT-ED than anticipated. It also implies that CBT-E affects interpersonal problem severity, despite it not explicitly targeting interpersonal problems. The latter could be explained by the focus CBT-E has on increasing interpersonal activities as part of increasing marginalised domains or as a means of avoiding binge eating between planned meals and snacks. Social isolation has been hypothesised to be a maintaining factor of eating disorder psychopathology (Murphy et al., 2012). By its focus on increasing interpersonal activities CBT-E may be addressing some of the interpersonal problems patients struggle with. An alternative explanation is that by directly addressing the patients' eating disorder psychopathology and associated behaviours CBT-E addresses a potential source of interpersonal conflict, and thus decreases the severity of their interpersonal problems. Thus, reduced interpersonal problem severity could by either a mediator of treatment outcome or an outcome of treatment in its own right.

This research replicates previous findings from two RCTs, one RCT (N=162) comparing IPT and CBT for binge eating disorder (BED) (Wilfley et al., 2002), and one sub-study in an RCT (N=50) comparing IPT for bulimia nervosa (IPT-BN) and CBT for bulimia nervosa (CBT-BN) (Fairburn, Jones, et al., 1993). These studies found that social adjustment and interpersonal problems improved in both treatments by the end of treatment. Only one RCT (N=220) has examined change in interpersonal problems during treatment. This RCT compared IPT-BN and CBT-BN and examined change in

interpersonal problems by week 10 of treatment (Wilson et al., 2002). No differences between treatments were found and no change in interpersonal problems was observed during this time. This suggests that change in interpersonal problems occurs sometime between week 10 and week 18/19 of treatment in IPT-ED. It also suggests that change in interpersonal problem severity has little to do with increasing interpersonal activities as part of managing urges to binge eat between planned meals and snacks, otherwise some change in interpersonal functioning might have been observed by this point. It may be that addressing eating disorder psychopathology improves interpersonal functioning in CBT-E, rather than addressing marginalised domains given that marginalised domains are not addressed in CBT-BN.

Whilst the interpersonal interview was carried out by independent raters, there would have been a natural discrepancy in terms of the amount of time spent identifying the interpersonal problems the interpersonal interview would have focussed on. In IPT-ED these would have been the problems that had been identified and agreed as targets of treatment between therapist and patient. In CBT-E these would have represented problems identified by therapist and patient on the basis of one conversation (although they may have represented problems brought up previously in other sessions). These problems would almost certainly have been explored in greater depth in IPT-ED, as that is the nature of the therapy. This may mean that the CBT-E problems were less accurately identified or perhaps that the problems specified were the simplest to identify and perhaps the simplest to address, which could explain why the two treatments apparently performed equivalently despite IPT-ED's greater focus on interpersonal problems.

The finding that IPT-ED appears to perform no better than CBT-E in addressing interpersonal problems supports the fact that interpersonal functioning has not been found to moderate treatment response in these two treatments. In fact, poor interpersonal functioning predicts poorer outcome irrespective of treatment, suggesting that despite directly targeting interpersonal problems, IPT does not perform better than CBT-E in addressing them (Z. Cooper et al., 2016; A. Jones, Lindekilde, Lübeck, & Clausen, 2015; Vall & Wade, 2015). The results of study 4a suggest that interpersonal functioning may explain why IPT-ED had poorer outcomes at end of treatment compared to CBT-E, although this needs to be interpreted with caution given that temporal ordering of change could not be established and no significant differences in interpersonal functioning was found between treatments.

Whilst further research is needed to ascertain the temporal ordering of change in interpersonal problems and eating disorder psychopathology, these results bring into question whether interpersonal problem severity is a specific mediator of IPT-ED or whether it is a shared mediator in the two treatments. The fact that both IPT-ED and CBT-E target, directly or indirectly, the same process (either as a potential mediator of outcome or as a result of improving eating disorder psychopathology), may explain why research examining IPT as a secondary treatment for non-responders to CBT has not had much success (e.g. Mitchell et al., 2002).

Future research should investigate change in interpersonal problem severity during the latter half treatment in order to assess the temporal ordering of change in interpersonal problem severity and eating disorder psychopathology. Such research should also attempt to investigate to what extent CBT-E indirectly addresses

interpersonal problems as part of treatment (i.e. the extent to which patients increase interpersonal activities as part of treatment) as well as to what extent improvements in eating disorder psychopathology predict engagement in interpersonal activities and reduced interpersonal conflict. A better understanding of the temporal ordering of these variables, and different aspects of interpersonal functioning would be helpful not only in determining whether interpersonal problem severity is a mediator of treatment outcomes, but also to gain a better understanding of the relative contribution of these variables in maintaining eating disorder psychopathology.

9.3. Methodological Considerations and Implications

9.3.1. Research Design

Embedding mediation studies within RCTs is an efficient and cost effective way of assessing purported mediators of treatment effects. The current studies were embedded within an RCT comparing CBT-E and IPT-ED, using block randomisation based on gender, eating disorder diagnosis and medication. Whilst neither therapists nor patients were blind to treatment allocation, assessments were carried by independent assessors who were. Treatment fidelity was good. Therefore, while the trial was subject to some risk of bias, this was minimised.

A major strength of the trial is that the purported mediators and the outcomes were assessed on a regular and frequent basis during treatment, at least in CBT-E. It also assessed several purported mediators of the core procedures of CBT-E and their specific outcomes in both treatments (e.g. binge eating, concern about shape / weight). This enabled a detailed analysis of these variables relationships over time in

both treatments. However, this was not the case for the purported mediator and outcome of IPT-ED, which was assessed less frequently in both treatments, and this limited the ability of the study to detect change in these variables in CBT-E and IPT-ED.

Another limitation of the study was the use of unvalidated, single item measures. As the majority of measures used in this study were either adapted or developed for use in this study little is known about their reliability. Whilst the outcome variables used in studies 2 and 3 were drawn from the Eating Disorder Examination (Fairburn, Cooper, & O'Connor, 2008), which is a validated measure, their phrasing and time span was adapted for use in this study, which might have affected their reliability. Furthermore, items used to assess purported mediators in studies 2 and 3 were added to the eating questionnaire, and did not derive from the previously validated measure. They also used a relatively narrow Likert scale range (1-5), which may have reduced the sensitivity of this scale. Whilst inter-rater reliability of the regular eating rating scale (RERS) binge eating rating scale (BERS), and interpersonal interview were assessed as part of this research, and was found to be acceptable, further research into their reliability would support the findings. To some extent, unreliability in these measures would have been controlled for due to inclusion of their baseline measurements, however, future research should consider using validated measures to ensure reliability of measures. Single item measures also tend to have poorer reliability than multi-item measures (Marsh, 1993). Using multi-item scales is doubly recommended when assessing the mediator construct (Iacobucci et al., 2007). This is because the mediator essentially lies on two paths within the model, being central to the indirect path from treatment to outcome (Iacobucci et al., 2007). Therefore, any measurement

error within the mediator will impact greatly upon multiple parts of the model (Iacobucci et al., 2007).

The study design could have been improved by randomising at the level of the purported mediator to improve causal inference regarding the effect of the purported mediator on the outcome. However such a design would be difficult in the context of a clinical RCT and be complicated by the presence of multiple mediators. An alternative would be to stratify randomisation based on the level of the purported mediator at baseline (e.g. by level of regularity of eating, interpersonal problem severity). However, in order to do this, careful consideration of potential cut-offs to help stratify scores would be needed, which would require further research. Stratification would also become increasingly complex as the number of purported mediators being studied increased. Another alternative would be the use of instrumental variables (Dunn et al., 2015; Emsley, Dunn, & White, 2010). These are variables that are related both to treatment and the purported mediator, but are unrelated to the outcome (except via its relationship to the purported mediator). As these variables are unrelated to the outcome directly, they can be used in the statistical analysis to approximate the effect of randomisation of the purported mediator, with one instrumental variable specified for each purported mediator. Such variables need to be considered carefully because, whilst instrumental variables can be constructed from any baseline variable, if the variable does not demonstrate strong relationships with both treatment and the purported mediator then using it as an instrument in the analysis may create more bias than the model would have without the inclusion of the instrument (Bound, Jaeger, & Baker, 1995).

As discussed in chapter 3, the use of statistical mediation analysis in the context of two active treatments is challenging, particularly when change is occurring simultaneously in both treatments. Whilst the research described in this thesis developed an analytic strategy designed to overcome some of these issues, RCTs comparing two active treatments are not optimal for assessing purported mediators of treatments. The endeavour to understand how psychological treatments work would be helped if future research could use an inactive control, such as a waitlist group. An alternative, to minimise time spent waiting for treatment, could be the use of a staged design using randomisation in which different cohorts of patients commence treatment at different times, based on different treatment stages (for example, at intervals of 4-5 weeks).

9.3.2. Statistical Analysis

The analytic strategy developed in this study took a model building approach to statistical mediation. It was designed to examine the relationships between treatment, the purported mediator and the outcome from different angles. It used multilevel modelling as a preliminary step (study a) to assess relevant non-specific predictors and moderators of treatment's effect on the purported mediators and the outcome, and for assessing trends over time in these variables. Structural equation modelling was then used to directly model the relationships between treatment, the purported mediator, and the outcome. This was done using three models, an autoregressive model (study b) to assess relationships between these variables at each time point, an alternative autoregressive model (study c) to test whether the purported mediator was better predicted by the outcome, and a latent growth curve

model (study d) to assess the relationship between change in the purported mediator and change in the outcome. This facilitated a detailed understanding of both how these variables related to each other during treatment, and what the overall effect of treatment and the purported mediator was in terms of change in the outcome.

Contemporary mediation analysts have argued that multilevel and structural equation modelling approaches are faulty as they don't accurately quantify the mediated effect (Dunn et al., 2015; Pearl, 2010; Rubin, 1974). Instead, they rely on examining the relationship between treatment, the purported mediator and the outcome. Whilst causal mediation analysis has been adapted to include longitudinal data, the approach is computationally intensive and potentially more sensitive to assumption violations compared to traditional methods (Taubman et al., 2009). There has also been little research with regards how this method performs in the context of two active treatments or conditions of treatment personalisation. Such a development would be helpful for the present type of research. An alternative approach would be to use multilevel modelling to select specific time points at which treatments could be compared in order to assess the contribution of a specific purported mediator at different key time points. For example, in study 1, change in regularity of eating and frequency of binge eating had occurred by week 2 of treatment in CBT-E but not in IPT-ED. An indirect effect could be calculated at this time point, and later on in treatment to compare mediated effects between treatments. Another option would be g-computation, although this approach is cumbersome with more complex models (Daniel et al., 2011).

Although purported mediators were investigated in isolation for the purpose of calculating more accurate indirect effects, in reality, the purported mediators assessed in CBT-E are likely to be interacting in some cases or sequential. How these different procedures fit together, in terms of interacting mediators or the individual contribution of different treatment procedures on overall outcome could be a focus of future research. However, such models would be extremely complex to assess statistically.

An intention to treat approach was taken to preserve the benefits gained from treatment randomisation (i.e. improved causal inference) and in order to estimate the extent to which a mediated effect might be expected to occur in clinical practice. This could have potentially diluted the mediated effect due to non-optimal treatment (e.g. having interrupted treatment from going on hold due to the need to treat a more acute condition such as depression) or not being the target group of the specific treatment procedure (e.g. patients who do not binge eat in the regular eating procedure would not have experienced a reduction in binge eating, although all patients receive this intervention). Such differences in treatment course may have affected the extent to which a purported mediator was activated during treatment. However, a per protocol analysis constrained to patients who adhered to the treatment protocol or to those subgroups targeted by the specific procedure (e.g. increasing regularity of eating to reduce frequency of binge eating in those who binge eat) would have further reduced an already small sample size and would have broken randomisation and have left the analysis open to confounding.

Due to lack of randomisation of the purported mediator, it is possible that the results were affected by unmeasured confounding. Whilst an instrumental variable approach was considered, as there were no planned instrumental variables and no natural instruments available, this approach was not used to avoid increasing bias. Another analytic option would have been the use of principal stratification (Dunn et al., 2015; Frangakis & Rubin, 2002; MacKinnon et al., 2007). This approach examines different patterns in how the purported mediator (e.g. high treatment compliance vs. low treatment compliance) relate to the outcome, treating the different patterns as strata to how treatment effects (direct and indirect) vary across the different patterns. Treatment effects can then be calculated across the different strata to get an average treatment effect, and to assess whether strata membership (in this case, high or low treatment compliance) is related to the outcome. Sensitivity analyses to test how robust the mediated effect was to potential violations of the assumption of unmeasured confounding would also have added to the reliability of the results (Imai, Keele, & Tingley, 2010).

Whilst the sample size (N=130) is respectable for an eating disorder treatment trial, it was relatively small for the complexity of the structural equation models run. Current rules of thumb indicate that a sample size of at least 200 is preferable, although a sample size of 100 can be sufficient in simple models (Kline, 2011). However, such rules apply to cross-sectional models. Thus far research into the sample size requirements of longitudinal structural equation modelling has been limited (Little, 2013). There has also been little research into the power to detect a mediated effect in these models (Cole & Maxwell, 2003). It is likely that tests of the mediated effect

were underpowered, which is a common problem in such analyses (MacKinnon et al., 2002).

A limitation of the study is that the indirect effect of treatment was calculated via all paths in the autoregressive model. This means that it reflects the impact that treatment had both via its direct effect on the outcome at earlier weeks in treatment, as well as its effect via the purported mediator. Results therefore need to be interpreted taking this into account.

Study b could have been improved by adding an additional step, which directly assessed to what extent modelling the relationship between the purported mediator and the outcome fit the data better than a model which only modelled the direct effects of treatment. Nested models such as this can be assessed statistically used the Chi-square. This could have helped clarify the results in studies 2b, 3b and 4b, where the relationship between the purported mediator and the outcome appeared to diminish over time. The stability of the autoregressive path might also have been better assessed using a similar approach in which the hypothesised (unconstrained autoregressive) model was compared to a model in which the autoregressive path coefficients were constrained to be equal. Wald tests are poorly powered (Little, 2013).

Although bootstrapping was used to address non-normality in the data, in a number of studies some of the bootstrap replications failed to converge. Failure rate was particularly high in the latent growth curve model, in which less than 40% of bootstrap replications failed to converge. This is indicative of some model instability, possibly related to extreme values (e.g. zero binge eating) in the dataset and small sample size.

Therefore, whilst results overall have been consistent with existing research, the results should be treated with caution. However, these extreme values are normative for the population studied, so it is likely they are representative.

Future research should consider examining the effect of sample size on power in autoregressive and latent growth curve models for statistical mediation. An exploration into how different trends over time in the purported mediator and the outcome affect the results could also be helpful (e.g. delayed response). Such research would inform interpretation, and provide evidence as to the robustness of these models. Simulation research assessing these models under conditions of delayed response and relapse between treatment weeks could be particularly helpful. Whilst more optimal measurement timing would be preferable, treatment week represents a meaningful measurement period, and requires less research resources than ecological momentary assessment.

9.3.2.1. Methods for Handling Treatment Personalisation

One of the strengths of this research is that it investigated the impact of treatment on purported mediators and outcomes in the context of treatment personalisation, where the timing of implementation of specific treatment procedures differed between patients (i.e. treatment was personalised). This is the first study to look at the impact of treatment on purported mediators and outcomes of psychological treatments in this context. Whilst previous research has examined the impact that the presence or absence, or dosage of different treatment components has on patient outcome (Dunn et al., 2012; Flach et al., 2015; Spangler et al., 2004), these studies have generally focused on broad treatment components, such as cognitive and / or behavioural

techniques (often grouped together), or structural aspects, such as completion of homework, rather than specific treatment procedures. For the most part, these studies have not examined the purported mediators associated with these treatment components, nor how change in the purported mediators relate to outcome (e.g. Dunn et al., 2012; Flach et al., 2015). Where research has examined purported mediators (i.e. Spangler et al., 2004), measurement of the timing of implementation of these components has been imprecise, based on rating to what extent cognitive or behavioural techniques in general have been used in the treatment session. Such research, at least in eating disorders, has also been limited by lack of use of a comparison condition.

Treatment personalisation was handled differently in the multilevel (study 3a) and structural equation (study 3b) models. In study a (multilevel models), personalisation of treatment was modelled as a piecewise trend, as an additional timeline increasing week on week from the point of implementation of the shape checking procedure. This effectively modelled what additional effect the shape checking procedure had on top of any other treatment effects. However, there were a number of potential coding options for modelling personalisation treatment (e.g. switch on/off, switch on). The method selected may not have been the best, however, it was the one that made the most clinical sense, given that decreasing concern about shape is hypothesised to be slow but cumulative. Nonetheless, further research into how best to model personalisation of treatment would be helpful to better inform decisions around coding. Similarly, how to model treatment personalisation in autoregressive structural equation models warrants further investigation. Inconsistency between studies may

be due to the methodological difficulty incorporating personalisation of treatment into the analyses. Future research should assess the accuracy of these models using simulated data that reflect possible clinical scenarios (e.g. mediation / no mediation).

9.4. General Discussion

The current studies focused on specific outcomes in CBT-E (e.g. concern about shape, frequency of binge eating), however, it is possible that these treatment procedures (and their purported mediators) affected other eating disorder symptoms as well (e.g. dietary restraint). The current research fails to address what these procedures contributed to addressing overall eating disorder psychopathology, although arguably both the specific and overall eating disorder outcomes are of interest. Better evidence regarding the effectiveness of individual treatment components may help persuade clinicians of the utility of individual treatment procedures. Streamlining treatments by removing their redundant components might also make them easier to implement. A better understanding of how active procedures work, including expected rate of change may help provide better guidance for clinicians (Z. Cooper & Bailey-Straebler, 2015; Shafran et al., 2009).

The extent to which these studies support the cognitive behavioural model as it is currently conceptualised needs to be considered. Over-evaluation of shape, weight, and eating is hypothesised to be the key maintaining factor that drives patients to both monitor and (attempt to) control their shape, weight and eating, typically through dieting. This can lead to weight loss, where the patient is undereating, and leave the patient vulnerable to binge eating due to physiological and psychological pressures to eat. Whilst the current research supports the hypothesised link between dieting and

binge eating, less clear is the role of monitoring in over-evaluation of shape, weight, and eating, and the relationship between over-evaluation of shape, weight, and eating, and dieting. Monitoring of shape and weight is hypothesised to reinforce over-evaluation of shape and weight, and it has therefore been suggested that frequent monitoring is a behaviour that needs to be addressed in treatment. There was little evidence for this being the case in these studies, with frequency of weight and shape checking being inconsistently related to concern about weight and shape during treatment. This may mean that the model is incorrect in this regard. Nonetheless these results need to be interpreted with caution due to the methodological limitations already mentioned, one of which includes small sample size and, consequently, low power. With regards to the relationship between over-evaluation of shape, weight, and eating, and dieting, this was not examined in any study. This is an area that will require attention in the future in order to fully assess whether the cognitive behavioural model functions as expected in the context of treatment.

This thesis examined only one purported mediator of IPT-ED; interpersonal problem severity. Whilst the central tenet of IPT is that interpersonal problems maintain eating disordered behaviours in both the short- and long-term, up until relatively recently there had been a lack of detail regarding the ways in which interpersonal problems maintained these behaviours. This may in part have been due to the more pragmatic origins of IPT in being developed as a comparator treatment for use in RCTs of CBT-BN (see chapter 1 for more detail). The lack of a specific model in IPT makes it difficult to conceptualise treatment processes and specify testable mediators. Whilst this study conceptualised IPT-ED as a cohesive, single unit of treatment, other authors have

contested this (e.g. Lipsitz & Markowitz, 2013). Different processes may be pertinent to different problem areas as they are approached in different ways. Furthermore, there may be multiple aspects underpinning interpersonal problems. Lipsitz & Markowitz (2013) propose that IPT addresses these interpersonal problems through four different processes. These are a) enhancing social support b) decreasing interpersonal stress c) facilitating emotional processing and d) improving interpersonal skills. Whilst further investigation into these processes might help to determine what factors are most relevant to decreasing interpersonal problems, the basic question as to whether addressing interpersonal problems improves eating disorder outcomes still needs resolving.

9.5. Conclusion

The research described in this thesis investigated three purported mediators of CBT-E and arguably the core purported mediator of IPT-ED. Only regularity of eating demonstrated a strong case for being a mediator of the effect of CBT-E (on frequency of binge eating). Overall, there is insufficient evidence to recommend treatment refinement on the basis of these results. In any event, the findings would require replication before adaptation to an established treatment was made. Nevertheless the findings provide direction for future research into these treatments. Further research is needed to identify the most appropriate methods for modelling this kind of data, taking into account likely scenarios that might be observed during treatment (e.g. temporary relapse, treatment personalisation). Embedding mediation studies within RCTs can provide better insight into different treatment processes. However, such studies require careful selection of measures and their timing of measurement.

Dismantling designs focusing on individual treatment procedures may be more suitable for assessing the causal effect of individual treatment procedures on treatment outcomes. Using a combination of these methods could represent the optimal approach to evaluating how treatment works, and the relative effects of different treatment procedures on the outcome.

References

Accurso, E. C., Wonderlich, S. A., Crosby, R. D., Smith, T. L., Klein, M. H., Mitchell, J. E., … Peterson, C. B. (2016). Predictors and moderators of treatment outcome in a randomized clinical trial for adults with symptoms of bulimia nervosa. *Journal of Consulting and Clinical Psychology*, *84*(2), 178–184. https://doi.org/10.1037/ccp0000073

Agras, W. S., Schneider, J. A., Arnow, B., Raeburn, S. D., & Telch, C. F. (1989a). Cognitive-behavioral and response-prevention treatments for bulimia nervosa. *Journal of Consulting and Clinical Psychology*, *57*(2), 215–221. https://doi.org/10.1037/0022-006X.57.2.215

Agras, W. S., Schneider, J. A., Arnow, B., Raeburn, S. D., & Telch, C. F. (1989b). Cognitive-behavioral and response-prevention treatments for bulimia nervosa. *Journal of Consulting and Clinical Psychology*, *57*(2), 215–221. https://doi.org/10.1037/0022-006X.57.2.215

Agras, W. S., Walsh, T., Fairburn, C. G., Wilson, G. T., & Kraemer, H. C. (2000). A multicenter comparison of cognitive-behavioral therapy and interpersonal psychotherapy for bulimia nervosa. *Archives of General Psychiatry*, *57*(5), 459–66.

American Psychiatric Association. (2006). *Practice guideline for the treatment of patients with eating disorders*. *The American Journal of Psychiatry* (3rd ed.). Washington, DC: Author.

Baron, R. M., & Kenny, D. A. (1986). The moderator-mediator variable distinction in social psychological research: conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology*, *51*(6), 1173–1182.

Bauer, D. J., Preacher, K. J., & Gil, K. M. (2006). Conceptualizing and testing random indirect effects and moderated mediation in multilevel models: New procedures and recommendations. *Psychological Methods*, *11*(2), 142–163. https://doi.org/10.1037/1082-989X.11.2.142

Beck, A. T. (1963). Thinking and depression I. Idiosyncratic content and cognitive distortions. *Archives of General Psychiatry*, *9*, 36–45.

Beck, A. T. (1964). Thinking and depression II. Theory and therapy. *Archives of General Psychiatry*, *10*, 561–571.

Beck, A. T. (1970). Cognitive therapy: Nature and relation to behavior therapy. *Behavior Therapy*, *1*(2), 184–200. https://doi.org/10.1016/S0005-7894(70)80030-2

Beck, A. T. (2005). The current state of cognitive therapy: A 40-year retrospective. *Achives of General Psychiatry*, *62*, 953–959.

Beck, A. T., Rush, A. J., Shaw, B., & Emery, G. (1984). *Cognitive Therapy of Depression*. New York, NY: Guilford Press.

Beck, A. T., Rush, A. J., Shaw, B. F., & Emery, G. (1979). *Cognitive Therapy of Depression.* New York, NY: Guilford Press.

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

Biesanz, J. C., Deeb-Sossa, N., Papadakis, A. A., Bollen, K. A., & Curran, P. J. (2004). The role of coding time in estimating and interpreting growth curve models. *Psychological Methods*, *9*(1), 30–52. https://doi.org/10.1037/1082-989X.9.1.30

Boomsma, A. (2000). Reporting analyses of covariance structures. *Structural Equation Modeling*, *7*(3), 461–483. https://doi.org/10.1207/S15328007SEM0703_6

Bound, J., Jaeger, D. A., & Baker, R. M. (1995). Problems with instrumental variables estimation when the correlation between the instruments and the endogenous explanatory variable is weak. *Journal of the American Statistical Association*, *90*(430), 443–450. https://doi.org/10.1080/01621459.1995.10476536

Brown, G. W., & Harris, T. O. (1978). *Social Origins of Depression: A Study of Psychiatric Disorder in Women.* London: Tavistock Publications.

Byrne, S. M., Fursland, A., Allen, K. L., & Watson, H. (2011). The effectiveness of enhanced cognitive behavioural therapy for eating disorders: An open trial. *Behaviour Research and Therapy*, *49*(4), 219–26. https://doi.org/10.1016/j.brat.2011.01.006

Calugi, S., Dalle Grave, R., Sartirana, M., & Fairburn, C. G. (2015). Time to restore body weight in adults and adolescents receiving cognitive behaviour therapy for anorexia nervosa. *Journal of Eating Disorders*, *3*(21). https://doi.org/10.1186/s40337-015-0057-z

Calugi, S., El Ghoch, M., & Dalle Grave, R. (2017). Body checking behaviors in anorexia nervosa. *International Journal of Eating Disorders*, *50*(4), 437–441. https://doi.org/10.1002/eat.22677

Calugi, S., Ghoch, M. El, & Grave, R. D. (2017). Body checking behaviors in anorexia nervosa. *International Journal of Eating Disorders*, *50*, 437–441. https://doi.org/10.1002/eat.22677

Carter, F. A., Jordan, J., McIntosh, V. V. W., Luty, S. E., McKenzie, J. M., Frampton, C. M. A., … Joyce, P. R. (2011). The long-term efficacy of three psychotherapies for anorexia nervosa: A randomized, controlled trial. *International Journal of Eating Disorders*, *44*(7), 647–654. https://doi.org/10.1002/eat.20879

Cheong, J., MacKinnon, D. P., & Khoo, S. T. (2003). Investigation of mediational processes using parallel process latent growth curve modeling. *Structural Equation Modeling*, *10*(2), 238–262. https://doi.org/10.1207/S15328007SEM1002_5

Cole, D. a, & Maxwell, S. E. (2003). Testing mediational models with longitudinal data: Questions and tips in the use of structural equation modeling. *Journal of Abnormal Psychology*, *112*(4), 558–577. https://doi.org/10.1037/0021-843X.112.4.558

Cooper, P. J., & Steere, J. (1995). A comparison of two psychological treatments for bulimia nervosa: Implications for models of maintenance. *Behaviour Research and Therapy*, *33*(8), 875–885. https://doi.org/10.1016/0005-7967(95)00033-T

Cooper, Z., Allen, E., Bailey-Straebler, S., Basden, S., Murphy, R., Connor, M. E. O., & Fairburn, C. G. (2016). Predictors and moderators of response to enhanced cognitive behaviour therapy and interpersonal psychotherapy for the treatment of eating disorders. *Behaviour Research and Therapy*, *84*, 9–13. https://doi.org/10.1016/j.brat.2016.07.002

Cooper, Z., & Bailey-Straebler, S. (2015). Disseminating evidence-based psychological treatments for eating disorders. *Current Psychiatry Reports*, *17*(3), 12. https://doi.org/10.1007/s11920-015-0551-7

Cooper, Z., & Fairburn, C. G. (2011). The evolution of "enhanced" cognitive behavior therapy for eating disorders: Learning from treatment nonresponse. *Cognitive and Behavioral Practice*, *18*(3), 394–402. https://doi.org/10.1016/j.cbpra.2010.07.007

Cuijpers, P., Andersson, G., Donker, T., & van Straten, A. (2011). Psychological treatment of depression: Results of a series of meta-analyses. *Nordic Journal of Psychiatry*, *65*(6), 354–364. https://doi.org/10.3109/08039488.2011.596570

Dalle Grave, R., Calugi, S., Conti, M., Doll, H., & Fairburn, C. G. (2013). Inpatient cognitive behaviour therapy for anorexia nervosa: A randomized controlled trial. *Psychotherapy and Psychosomatics*, *82*(6), 390–398. https://doi.org/10.1159/000350058

Dalle Grave, R., Calugi, S., Doll, H. A., & Fairburn, C. G. (2013). Enhanced cognitive behaviour therapy for adolescents with anorexia nervosa: An alternative to family therapy? *Behaviour Research and Therapy*, *51*(1), R9–R12. https://doi.org/10.1016/j.brat.2012.09.008

Dalle Grave, R., Calugi, S., El Ghoch, M., Conti, M., & Fairburn, C. G. (2014). Inpatient cognitive behavior therapy for adolescents with anorexia nervosa: Immediate and longer-term effects. *Frontiers in Psychiatry*, *5*, 1–7. https://doi.org/10.3389/fpsyt.2014.00014

Dalle Grave, R., Calugi, S., Sartirana, M., & Fairburn, C. G. (2015). Transdiagnostic cognitive behaviour therapy for adolescents with an eating disorder who are not underweight. *Behaviour Research and Therapy*, *73*, 79–82. https://doi.org/10.1016/j.brat.2015.07.014

Daniel, R. M., De Stavola, B., & Cousens, S. (2011). gformula: Estimating causal effects in the presence of time-dependent confounding or mediation using the g-computation formula. *The Stata Journal*, *11*(4), 479–517.

Daniel, R. M., De Stavola, B. L., Cousens, S. N., & Vansteelandt, S. (2015). Causal mediation analysis with multiple mediators. *Biometrics*, *71*(1), 1–14. https://doi.org/10.1111/biom.12248

Davies, S. (2014). *Annual report of the chief medical officer 2014. The health of the 51%: Women*. Department of Health.

De Young, K. P., Lavender, J. M., Crosby, R. D., Wonderlich, S. A., Engel, S. G., Mitchell, J. E., … Le Grange, D. (2014). Bidirectional associations between binge eating and restriction in anorexia nervosa. An ecological momentary assessment study. *Appetite*, *83*, 69–74. https://doi.org/10.1016/j.appet.2014.08.014

Doss, B. D., & Texas, A. (2006). Investigating treatment mediators when simple random assignment to a control group is not possible. *Clinical Psychology: Science and Practice*, *13*(4), 321–336.

Driessen, E., & Hollon, S. D. (2010). Cognitive behavioral therapy for mood disorders: Efficacy,

moderators and mediators. *Psychiatric Clinics of North America*, *33*(3), 537–555. https://doi.org/10.1016/j.psc.2010.04.005

Dunn, G., Emsley, R., Liu, H., Landau, S., Green, J., White, I., & Pickles, A. (2015). Evaluation and validation of social and psychological markers in randomised trials of complex interventions in mental health: A methodological research programme. *Health Technology Assessment*, *19*(93), 1–116. https://doi.org/10.3310/hta19930

Dunn, G., Fowler, D., Rollinson, R., Freeman, D., Kuipers, E., Smith, B., … Bebbington, P. (2012). Effective elements of cognitive behaviour therapy for psychosis: Results of a novel type of subgroup analysis based on principal stratification. *Psychological Medicine*, *42*, 1057–1068. https://doi.org/10.1017/S0033291711001954

Elkin, I., Parloff, M. B., Hadley, S. W., & Autry, J. H. (1985). NIMH Treatment of Depression Collaborative Research Program. Background and research plan. *Archives of General Psychiatry*, *42*(3), 305–316. https://doi.org/10.1001/archpsyc.1985.01790260103013

Elkin, I., Shea, M. T., Watkins, J. T., Imber, S. D., Sotsky, S. M., Collins, J. F., … Parloff, M. B. (1989). National Institute of Mental Health Treatment of Depression Collaborative Research Program. General effectiveness of treatments. *Archives of General Psychiatry*, *46*(11), 971–982. https://doi.org/10.1001/archpsyc.1989.01810110013002

Ellison, J. M., Simonich, H. K., Wonderlich, S. A., Crosby, R. D., Cao, L., Mitchell, J. E., … Peterson, C. B. (2016). Meal patterning in the treatment of bulimia nervosa. *Eating Behaviors*, *20*, 39–42. https://doi.org/10.1016/j.eatbeh.2015.11.008

Emsley, R., Dunn, G., & White, I. R. (2010). Mediation and moderation of treatment effects in randomised controlled trials of complex interventions. *Statistical Methods in Medical Research*, *19*(3), 237–270. https://doi.org/10.1177/0962280209105014

Fairburn, C. G. (1981). A cognitive behavioural approach to the treatment of bulimia. *Psychological Medicine*, *11*(4), 707–711.

Fairburn, C. G. (1984). Cognitive-behavioral treatment for bulimia. In D. M. Garner & P. E. Garfinkel (Eds.), *Handbook of Psychotherapy for Anorexia Nervosa and Bulimia Nervosa* (pp. 160–192). New York, NY: Guilford Press.

Fairburn, C. G. (1992). Interpersonal psychotherapy for bulimia nervosa. In G. L. Klerman & M. M. Weissman (Eds.), *New Applications of Interpersonal Psychotherapy* (pp. 353–378). Washington DC: American Psychiatric Press.

Fairburn, C. G. (1995). *Overcoming Binge Eating*. New York, NY: Guilford Press.

Fairburn, C. G. (1997). Interpersonal psychotherapy for bulimia nervosa. In D. M. Garner & P. E. Garfinke (Eds.), *Handbook of Treatment for Eating Disorders* (pp. 278–294). New York, NY: Guilford Press.

Fairburn, C. G. (2008). *Cognitive Behavior Therapy and Eating Disorders*. New York, NY: Guilford Press.

Fairburn, C. G., Bailey-Straebler, S., Basden, S., Doll, H. A., Jones, R., Murphy, R., … Zafra Cooper. (2015). A transdiagnostic comparison of enhanced cognitive behaviour therapy (CBT-E) and interpersonal psychotherapy in the treatment of eating disorders. *Behaviour Research and Therapy*, *70*, 64–71. https://doi.org/10.1016/j.brat.2015.04.010

Fairburn, C. G., & Beglin, S. J. (2008). Eating Disorder Examination Questionnaire (EDE-Q 6.0). In C. G. Fairburn (Ed.), *Cognitive Behavior Therapy and Eating Disorders* (pp. 309–313). New York, NY: Guilford Press.

Fairburn, C. G., & Cooper, Z. (2011). Eating disorders, DSM-5 and clinical reality. *British Journal of Psychiatry*, *198*(1), 8–10. https://doi.org/10.1192/bjp.bp.110.083881

Fairburn, C. G., Cooper, Z., & Cooper, P. J. (1986). The clinical features and maintenance of bulimia nervosa. In K. D. Brownell & J. P. Foreyt (Eds.), *Handbook of Eating Disorders: Physiology, Psychology, and Treatment of Obesity, Anorexia and Bulimia* (pp. 289–404). New York, NY: Basic Books.

Fairburn, C. G., Cooper, Z., Doll, H. A., O'Connor, M. E., Bohn, K., Hawker, D. M., … Palmer, R. L. (2009). Transdiagnostic cognitive-behavioral therapy for patients with eating disorders: A two-site trial with 60-week follow-up. *American Journal of Psychiatry*, *166*(3), 311–319. https://doi.org/10.1176/appi.ajp.2008.08040608

Fairburn, C. G., Cooper, Z., Doll, H. A., O'Connor, M. E., Palmer, R. L., & Dalle Grave, R. (2013). Enhanced cognitive behaviour therapy for adults with anorexia nervosa: A UK–Italy study. *Behaviour Research and Therapy*, *51*(1), R2–R8. https://doi.org/10.1016/j.brat.2012.09.010

Fairburn, C. G., Cooper, Z., & O'Connor, M. (2008). Eating Disorder Examination (edition 16.0D). In C. G.

Fairburn (Ed.), *Cognitive Behavior Therapy and Eating Disorders* (pp. 265–308). New York, NY: Guilford Press.
Fairburn, C. G., Cooper, Z., & Shafran, R. (2003). Cognitive behaviour therapy for eating disorders: a "transdiagnostic" theory and treatment. *Behaviour Research and Therapy*, *41*(5), 509–528. https://doi.org/10.1016/S0005-7967(02)00088-8
Fairburn, C. G., Cooper, Z., Shafran, R., Bohn, K., & Hawker, D. (2008a). Clinical perfectionism, core low self-esteem, and interpersonal problems. In C. G. Fairburn (Ed.), *Cognitive Behavior Therapy and Eating Disorders* (pp. 197–230). New York, NY: Guilford Press.
Fairburn, C. G., Cooper, Z., Shafran, R., Bohn, K., Hawker, D., Murphy, R., & Straebler, S. (2008b). Enhanced cognitive behavior therapy for eating disorders: The core protocol. In C. G. Fairburn (Ed.), *Cognitive Behavior Therapy and Eating Disorders* (pp. 47–193). New York, NY: Guilford Press.
Fairburn, C. G., & Harrison, P. J. (2003). Eating disorders. *The Lancet*, *361*, 407–416.
Fairburn, C. G., Jones, R., Peveler, R. C., Carr, S. J., Solomon, R. A., O'Connor, M. E., … Hope, R. A. (1991). Three psychological treatments for bulimia nervosa. A comparative trial. *Archives of General Psychiatry*, *48*(5), 463–469. https://doi.org/10.1001/archpsyc.1991.01810290075014
Fairburn, C. G., Jones, R., Peveler, R. C., Hope, R. A., & O'Connor, M. (1993). Psychotherapy and bulimia nervosa: Longer-term effects of interpersonal psychotherapy, behavior therapy, and cognitive behavior therapy. *Archives of General Psychiatry*, *50*(6), 419–428. https://doi.org/10.1001/archpsyc.1993.01820180009001
Fairburn, C. G., Kirk, J., O'Connor, M., & Cooper, P. J. (1986). A comparison of two psychological treatments for bulimia nervosa. *Behaviour Research and Therapy*, *24*(6), 629–643. https://doi.org/10.1016/0005-7967(86)90058-6
Fairburn, C. G., Marcus, M. D., & Wilson, G. T. (1993). Cognitive-behavioral therapy for binge eating and bulimia nervosa: A comprehensive treatment manual. In C. G. Fairburn & G. T. Wilson (Eds.), *Binge-Eating. Nature, Assessment and Treatment* (pp. 361–404). New York, NY.
Fairchild, A. J., & MacKinnon, D. P. (2009). A general model for testing mediation and moderation effects. *Prevention Science*, *10*(2), 87–99. https://doi.org/10.1007/s11121-008-0109-6
Fan, X., & Sivo, S. A. (2005). Sensitivity of fit indexes to misspecified structural or measurement model components: Rationale of two-index strategy revisited. *Structural Equation Modeling: A Multidisciplinary Journal*, *12*(3), 343–367. https://doi.org/10.1207/s15328007sem1203_1
Feinian, C., Curran, P. J., Bollen, K. A., Kirby, J., & Paxton, P. (2008). An empirical evaluation of the use of fixed cutoff points in RMSEA test statistic in structural equation models. *Sociological Methods & Research*, *36*(4), 462–494. https://doi.org/10.1177/0049124108314720
Flach, C., French, P., Dunn, G., Fowler, D., Gumley, A. I., Birchwood, M., … Morrison, A. P. (2015). Components of therapy as mechanisms of change in cognitive therapy for people at risk of psychosis: Analysis of the EDIE-2 trial. *British Journal of Psychiatry*, *207*(2), 123–129. https://doi.org/10.1192/bjp.bp.114.153320
Frangakis, C. E., & Rubin, D. B. (2002). Principal stratification in causal inference. *Biometrics*, *58*(1), 21–29. https://doi.org/10.2307/3068286
Freeman, C. P. L., Barry, F., Dunkeld-Turnbull, J., & Henderson, A. (1988). Controlled trial of psychotherapy for bulimia nervosa. *British Medical Journal*, *296*(6621), 521–525.
Fritz, M. S., Cox, M. G., & Mackinnon, D. P. (2013). Increasing statistical power in mediation models without increasing sample size. *Evaluation & the Health Professions*, *38*(3), 343–366. https://doi.org/10.1177/0163278713514250
Fritz, M. S., & Mackinnon, D. P. (2007). Required sample size to detect the mediated effect. *Psychological Science*, *18*(3), 233–9. https://doi.org/10.1111/j.1467-9280.2007.01882.x
Greenland, S., Pearl, J., & Robins, J. M. (1999). Causal diagrams for epidemiologic research. *Epidemiology (Cambridge, Mass.)*, *10*(1), 37–48. https://doi.org/10.1097/00001648-199901000-00008
Hay, P. J. (2013). A systematic review of evidence for psychological treatments in eating disorders: 2005-2012. *The International Journal of Eating Disorders*, *46*(5), 462–9. https://doi.org/10.1002/eat.22103
Hay, P. J., Bacaltchuk, J., Stefano, S., & Kashyap, P. (2009). Psychological treatments for bulimia nervosa and binging (review). *Cochrane Database of Systematic Reviews*, (4). https://doi.org/10.1002/14651858.CD000562.pub3
Hay, P. J., Chinn, D., Forbes, D., Madden, S., Newton, R., Sugenor, L., … Ward, W. (2014). Royal Australian and New Zealand College of Psychiatrists clinical practice guidelines for the treatment

of eating disorders. *Australian & New Zealand Journal of Psychiatry*, *48*(11), 977–1008. https://doi.org/10.1177/0004867414555814

Hertzog, C., & Nesselroade, J. R. (1987). Beyond autoregressive models: Some implications of the trait-state distinction for the structural modeling of developmental change. *Child Development*, *58*(1), 93–109.

Hilbert, A., Bishop, M. E., Stein, R. I., Tanofsky-Kraff, M., Swenson, A. K., Welch, R. R., & Wilfley, D. E. (2012). Long-term efficacy of psychological treatments for binge eating disorder. *The British Journal of Psychiatry*, *200*(3), 232–237. https://doi.org/10.1192/bjp.bp.110.089664

Hilbert, A., Saelens, B. E., Stein, R. I., Mockus, D. S., Welch, R. R., Matt, G. E., & Wilfley, D. E. (2007). Pretreatment and process predictors of outcome in interpersonal and cognitive behavioral psychotherapy for binge eating disorder. *Journal of Consulting and Clinical Psychology*, *75*(4), 645–651. https://doi.org/10.1037/0022-006X.75.4.645

Hollon, S., DeRubeis, R., & Evans, M. (1987). Causal mediation of change in treatment for depression: Discriminating between nonspecificity and noncausality. *Psychological Bulletin*, *102*(1), 139–149. https://doi.org/10.1037/0033-2909.102.1.139

Holmbeck, G. N. (1997). Toward terminological, conceptual, and statistical clarity in the study of mediators and moderators: Examples from the child-clinical and pediatric psychology literatures. *Journal of Consulting and Clinical Psychology*, *65*(4), 599–610.

Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal*, *6*(1), 1–55. https://doi.org/10.1080/10705519909540118

Iacobucci, D., Saldanha, N., & Deng, X. (2007). A meditation on mediation: Evidence that structural equations models perform better than regressions. *Journal of Consumer Psychology*, *17*(2), 140–154. https://doi.org/10.1016/S1057-7408(07)70020-7

Iacovino, J. M., Gredysa, D. M., Altman, M., & Wilfley, D. E. (2012). Psychological treatments for binge eating disorder. *Current Psychiatry Reports*, *14*(4), 432–446. https://doi.org/10.1007/s11920-012-0277-8

Imai, K., Keele, L., & Tingley, D. (2010). A general approach to causal mediation analysis. *Psychological Methods*, *15*(4), 309–34. https://doi.org/10.1037/a0020761

James, L. R., Mulaik, S. A., & Brett, J. M. (2006). A tale of two methods. *Organizational Research Methods*, *9*(2), 233–244. https://doi.org/10.1177/1094428105285144

Jones, A., Lindekilde, N., Lübeck, M., & Clausen, L. (2015). The association between interpersonal problems and treatment outcome in the eating disorders: A systematic review. *Nordic Journal of Psychiatry*, *9488*(8), 563–573. https://doi.org/10.3109/08039488.2015.1019924

Jones, R., Peveler, R. C., Hope, R. A., & Fairburn, C. G. (1993). Changes during treatment for bulimia nervosa: A comparison of three psychological treatments. *Behaviour Research and Therapy*, *31*(5), 479–485. https://doi.org/10.1016/0005-7967(93)90128-H

Judd, C. M., & Kenny, D. A. (1981). Process anlaysis: Estimating mediation in treatement evaluations. *Evaluation Review*, *5*(5), 602–619.

Kachani, A. T., Brasiliano, S., Cordás, T. A., & Hochgraf, P. B. (2013). Body checking and associated cognitions among Brazilian outpatients with eating disorders and nonpsychiatric controls. *Body Image*, *10*(1), 127–130. https://doi.org/10.1016/j.bodyim.2012.10.006

Kazdin, A. E. (2007). Mediators and mechanisms of change in psychotherapy research. *Annual Review of Clinical Psychology*, *3*, 1–27.

Kazdin, A. E., & Nock, M. K. (2003). Delineating mechanisms of change in child and adolescent therapy: Methodological issues and research recommendations. *Journal of Child Psychology and Psychiatry*, *44*(8), 1116–1129.

Kazdin, A. E., & Weisz, J. R. (1998). Identifying and developing empirically supported child and adolescent treatments. *Journal of Consulting and Clinical Psychology*, *66*(1), 19–36. https://doi.org/10.1037/0022-006X.66.1.19

Keele, L. (2015). Causal mediation analysis: Warning! Assumptions ahead. *American Journal of Evaluation*, *36*(4), 500–513. https://doi.org/10.1177/1098214015594689

Kenny, D. A., Korchmaros, J. D., & Bolger, N. (2003). Lower level mediation in multilevel models. *Psychological Methods*, *8*(2), 115–128. https://doi.org/10.1037/1082-989X.8.2.115

Kenny, D. A., & McCoach, D. B. (2003). Effect of the number of variables on measures of fit in structural equation modeling. *Structural Equation Modeling: A Multidisciplinary Journal*, *10*(3), 333–351.

https://doi.org/10.1207/S15328007SEM1003_1
Keys, A., Brozek, J., Henshel, A., Mickelson, O., & Taylor, H. L. (1950). *The Biology of Human Starvation* (Vols. 1-2). Minneapolis, MN: University of Minnesota Press.
Kirkley, B. G., Schneider, J. A., Agras, W. S., & Bachman, J. A. (1985). Comparison of two group treatments for bulimia. *Journal of Consulting and Clinical Psychology*, *53*(1), 43–48. https://doi.org/10.1037/0022-006X.53.1.43
Klerman, G. L., Weissman, M. M., Rounsaville, B. J., & Chevron, E. S. (1984). *Interpersonal Psychotherapy for Depression.* New York, NY: Basic Books. Retrieved from http://www.amazon.co.uk/Interpersonal-Psychotherapy-Depression-Gerald-Klerman/dp/0465033962/ref=sr_1_1?s=books&ie=UTF8&qid=1397049247&sr=1-1&keywords=interpersonal+psychotherapy+depression
Kline, R. B. (2011). *Principles and Practice of Structural Equation Modelling* (3rd ed.). New York, NY, NY: Guilford Press.
Kline, R. B. (2015). The mediation myth. *Basic and Applied Social Psychology*, *37*(4), 202–213. https://doi.org/10.1080/01973533.2015.1049349
Knott, S., Woodward, D., Hoefkens, A., & Limbert, C. (2015). Cognitive behaviour therapy for bulimia nervosa and eating disorders not otherwise specified: Translation from randomized controlled trial to a clinical setting. *Behavioural and Cognitive Psychotherapy*, *43*(6), 641–654. https://doi.org/10.1017/S1352465814000393
Kraemer, H. C. (2001). How do risk factors work together? Mediators, moderators, and independent, overlapping, and proxy risk factors. *American Journal of Psychiatry*, *158*(6), 848–856. https://doi.org/10.1176/appi.ajp.158.987.848
Kraemer, H. C. (2008). Toward non-parametric and clinically meaningful moderators and mediators. *Statistics in Medicine*, *27*(10), 1679–92. https://doi.org/10.1002/sim.3149
Kraemer, H. C. (2010). Moderators and mediators: The MacArthur updated view. In A. Steptoe (Ed.), *Handbook of Behavioral Medicine* (pp. 869–880). New York, NY: Springer. https://doi.org/10.1007/978-0-387-09488-5_55
Kraemer, H. C., & Gibbons, R. D. (2009). Why does the randomized clinical trial methodology so often mislead clinical decision making? *Psychiatric Annals*, *39*(7), 736–745. https://doi.org/10.3928/00485713-20090625-06
Kraemer, H. C., Kiernan, M., Essex, M., & Kupfer, D. J. (2008). How and why criteria defining moderators and mediators differ between the Baron & Kenny and MacArthur approaches. *Health Psychology*, *27*(2 Suppl), S101–S108. https://doi.org/10.1037/0278-6133.27.2(Suppl.).S101
Kraemer, H. C., Wilson, G. T., Fairburn, C. G., & Agras, W. S. (2002). Mediators and moderators of treatment effects in randomized clinical trials. *Archives of General Psychiatry*, *59*(10), 877. https://doi.org/10.1001/archpsyc.59.10.877
Krull, J. L., & MacKinnon, D. P. (1999). Multilevel mediation modeling in group-based intervention studies. *Evaluation Review*, *23*(4), 418–444. https://doi.org/10.1177/0193841X9902300404
Krull, J. L., & MacKinnon, D. P. (2001). Multilevel modeling of individual and group level mediated effects. *Multivariate Behavioral Research*, *36*(2), 249–277. https://doi.org/10.1207/S15327906MBR3602_06
Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics*, *33*(1), 159–174.
Legenbauer, T., Schütt-Strömel, S., Hiller, W., & Vocks, S. (2011). Predictors of improved eating behaviour following body image therapy: A pilot study. *European Eating Disorders Review*, *19*(2), 129–137. https://doi.org/10.1002/erv.1017
Leitenberg, H., & Rosen, J. (1989). Cognitive-behavioral therapy with and without exposure plus response prevention in treatment of bulimia nervosa: Comment on Agras, Schneider, Arnow, Raeburn, and Telch. *Journal of Consulting and Clinical Psychology*, *57*(6), 776–777. Retrieved from http://www.scopus.com/inward/record.url?eid=2-s2.0-0024788120&partnerID=40&md5=de464e863488deb732d7b6dbcb48d19c
Lemmens, L. H. J. M., Arntz, A., Peeters, F., Hollon, S. D., Roefs, A., & Huibers, M. J. H. (2015). Clinical effectiveness of cognitive therapy v. interpersonal psychotherapy for depression: Results of a randomized controlled trial. *Psychological Medicine*, *45*(10), 2095–2110. https://doi.org/10.1017/S0033291715000033
Linardon, J., Brennan, L., & de la Piedad Garcia, X. (2016). Rapid response to eating disorder treatment:

A systematic review and meta-analysis. *International Journal of Eating Disorders*, *49*(10), 905–919. https://doi.org/10.1002/eat.22595

Linardon, J., de la Piedad Garcia, X., & Brennan, L. (2017). Predictors, moderators, and mediators of treatment outcome following manualised cognitive-behavioural therapy for eating disorders: A systematic review. *European Eating Disorders Review*, *25*(1), 3–12. https://doi.org/10.1002/erv.2492

Lipsitz, J. D., & Markowitz, J. C. (2013). Mechanisms of change in interpersonal therapy (IPT). *Clinical Psychology Review*, *33*(8), 1134–1147. https://doi.org/10.1016/j.cpr.2013.09.002

Little, T. D. (2013). *Longitudinal Structural Equation Modelling*. New York, NY, NY: Guilford Press.

Little, T. D., Cunningham, W. A., Shahar, G., & Widaman, K. F. (2002). To parcel or not to parcel: Exploring the question, weighing the merits. *Structural Equation Modeling*, *9*(2), 151–173.

Little, T. D., Rhemtulla, M., Gibson, K., & Schoemann, A. M. (2013). Why the items versus parcels controversy needn't be one. *Psychological Methods*, *18*(3), 285–300. https://doi.org/10.1037/a0033266

Loeb, K. L., Wilson, G. T., Labouvie, E., Pratt, E. M., Hayaki, J., Walsh, B. T., … Fairburn, C. G. (2005). Therapeutic alliance and treatment adherence in two interventions for bulimia nervosa: A study of process and outcome. *Journal of Consulting and Clinical Psychology*, *73*(6), 1097–1106. https://doi.org/10.1037/0022-006X.73.6.1097

Luty, S. E., Carter, J. D., McKenzie, J. M., Rae, A. M., Frampton, C. M. A., Mulder, R. T., & Joyce, P. R. (2007). Randomised controlled trial of interpersonal psychotherapy and cognitive-behavioural therapy for depression. *The British Journal of Psychiatry*, *190*(6), 496–502. https://doi.org/10.1192/bjp.bp.106.024729

MacKinnon, D. P. (2008). *Introduction to Statistical Mediation Analysis*. New York, NY: Taylor & Francis Group.

MacKinnon, D. P., Fairchild, A. J., & Fritz, M. S. (2007). Mediation analysis. *Annual Review of Psychology*, *58*(1), 593–614. https://doi.org/10.1146/annurev.psych.58.110405.085542

MacKinnon, D. P., Krull, J. L., & Lockwood, C. M. (2000). Equivalence of the mediation, confounding and suppression effect. *Prevention Science*, *1*(4), 173–181. https://doi.org/10.1023/A1026595011371

MacKinnon, D. P., Lockwood, C. M., Hoffman, J. M., West, S. G., & Sheets, V. (2002). A comparison of methods to test mediation and other intervening variable effects. *Psychological Methods*, *7*(1), 83–104. https://doi.org/10.1037//1082-989X.7.1.83

Markowitz, J. C., & Weissman, M. M. (2012). Interpersonal psychotherapy: Past, present and future. *Clinical Psychology & Psychotherapy*, *19*(2), 99–105. https://doi.org/10.1002/cpp.1774

Marsh, H. W. (1993). Stability of individual differences in multiwave panel studies: Comparison of simplex models and one-factor models. *Journal of Educational Measurement*, *30*(2), 157–183.

Marsh, H. W., Hau, K.-T., & Wen, Z. (2004). In search of golden rules: Comment on hypothesis-testing approaches to setting cutoff values for fit indexes and dangers in overgeneralizing Hu and Bentler's (1999) findings. *Structural Equation Modeling: A Multidisciplinary Journal*, *11*(3), 320–341. https://doi.org/10.1207/s15328007sem1103_2

McCoach, D. B., & Kaniskan, B. (2010). Using time-varying covariates in multilevel growth models. *Frontiers in Psychology*, *1*(17), 1–12. https://doi.org/10.3389/fpsyg.2010.00017

McIntosh, V. V. W., Jordan, J., Carter, F. A., Luty, S. E., McKenzie, J. M., Bulik, C. M., … Joyce, P. R. (2005). Three psychotherapies for anorexia nervosa: A randomized, controlled trial. *American Journal of Psychiatry*, *162*(4), 741–747.

Mintz, L. B., Awad, G. H., Stinson, R. D., Bledman, R. A., Coker, A. D., Kashubeck-West, S., & Connelly, K. (2013). Weighing and body monitoring among college women: The scale number as an emotional barometer. *Journal of College Student Psychotherapy*, *27*(1), 78–91. https://doi.org/10.1080/87568225.2013.739039

Mitchell, J. E., Halmi, K., Wilson, G. T., Agras, W. S., Kraemer, H., & Crow, S. (2002). A randomized secondary treatment study of women with bulimia nervosa who fail to respond to CBT. *International Journal of Eating Disorders*, *32*(3), 271–281. https://doi.org/10.1002/eat.10092

Murphy, R., Cooper, Z., Hollon, S. D., & Fairburn, C. G. (2009). How do psychological treatments work? Investigating mediators of change. *Behaviour Research and Therapy*, *47*(1), 1–5. https://doi.org/10.1016/j.brat.2008.10.001

Murphy, R., Straebler, S., Basden, S., Cooper, Z., & Fairburn, C. G. (2012). Interpersonal psychotherapy for eating disorders. *Clinical Psychology and Psychotherapy*, *19*(2), 150–158.

https://doi.org/10.1002/cpp.1780
Muthen, B., & Curran, P. J. (1997). General longitudinal modeling of individual differences in experimental designs: A latent variable framework for analysis and power estimation. *Psychological Methods*, *2*(4), 371–402. Retrieved from http://psycnet.apa.org/journals/met/2/4/371/
National Collaborating Centre for Mental Health. (2004). *Eating Disorders: Core Interventions in the Treatment and Management of Anorexia Nervosa, Bulimia Nervosa and Related Eating Disorders (CG9)*. London: The British Psychological Society and Gaskell. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/23346610
National Institute for Health and Care Excellence. (2004). *Eating disorders. Core Interventions in the Treatment and Management of Anorexia Nervosa, Bulimia Nervosa, and Related Eating Disorders (CG9)*. London: National Institute for Health and Care Excellence.
National Institute for Health and Care Excellence. (2009). *Depression in adults: Recognition and Management (CG90)*. London: Author.
National Institute for Health Research. (2017). Efficacy and Mechanism Evaluation. Retrieved May 6, 2017, from http://www.nihr.ac.uk/funding-and-support/funding-for-research-studies/funding-programmes/efficacy-and-mechanism-evaluation/
Nazar, B. P., Gregor, L. K., Albano, G., Marchica, A., Lo Coco, G., Cardi, V., & Treasure, J. (2017). Early response to treatment in eating disorders: A systematic review and a diagnostic test accuracy meta-analysis. *European Eating Disorders Review*, *25*(2), 67–79. https://doi.org/10.1002/erv.2495
Onslow, L., Woodward, D., Hoefkens, T., & Waddington, L. (2016). Experiences of enhanced cognitive behaviour therapy for bulimia nervosa. *Behavioural and Cognitive Psychotherapy*, *44*(2), 168–178. https://doi.org/10.1017/S135246581400068X
Paykel, E., Myers, J., Dienelt, M., Klerman, G., Lindenthal, J., & Pepper, M. (1969). Life events and depression: A controlled study. *Archives of General Psychiatry*, *21*(6), 753–760.
Peeters, F., Huibers, M., Roelofs, J., van Breukelen, G., Hollon, S. D., Markowitz, J. C., … Arntz, A. (2013). The clinical effectiveness of evidence-based interventions for depression: A pragmatic trial in routine practice. *Journal of Affective Disorders*, *145*(3), 349–355. https://doi.org/10.1016/j.jad.2012.08.022
Polnay, A., James, V. A. W., Hodges, L., Murray, G. D., Munro, C., & Lawrie, S. M. (2014). Group therapy for people with bulimia nervosa: Systematic review and meta-analysis. *Psychological Medicine*, *44*(11), 2241–2254. https://doi.org/10.1017/S0033291713002791
Poulsen, S., Lunn, S., Daniel, S. I. F., Folke, S., Mathiesen, B. B., Katznelson, H., & Fairburn, C. G. (2014). A randomized controlled trial of psychoanalytic psychotherapy or cognitive-behavioral therapy for bulimia nervosa. *The American Journal of Psychiatry*, *171*(1), 109–116. https://doi.org/10.1176/appi.ajp.2013.12121511
Preacher, K. J. (2015). Advances in mediation analysis: A survey and synthesis of new developments. *Annual Review of Psychology*, *66*(1), 825–852. https://doi.org/10.1146/annurev-psych-010814-015258
Rigdon, E. E. (1996). CFI versus RMSEA: A comparison of two fit indexes for structural equation modeling. *Structural Equation Modeling: A Multidisciplinary Journal*, *3*(4), 369–379. https://doi.org/10.1080/10705519609540052
Rosenberg, M. (1965). *Society and the adolescent self-image*. Princeton, USA: Princeton University Press.
Rubin, D. B. (1974). Estimating causal effects of treatments in randomized and nonrandomized studies. *Journal of Educational Psychology*, *66*(5), 688–701. https://doi.org/10.1037/h0037350
Russell, G. (1979). Bulimia nervosa: An ominous variant of anorexia nervosa. *Psychological Medicine*, *9*(3), 429–448. https://doi.org/10.1017/S0033291700031974
Selig, J. P., & Preacher, K. J. (2009). Mediation models for longitudinal data in developmental research. *Research in Human Development*, *6*(2–3), 144–164. https://doi.org/10.1080/15427600902911247
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*(11), 902–909. https://doi.org/10.1016/j.brat.2009.07.003
Shafran, R., Fairburn, C. G., Robinson, P., & Lask, B. (2004). Body checking and its avoidance in eating disorders. *International Journal of Eating Disorders*, *35*(1), 93–101. https://doi.org/10.1002/eat.10228
Shah, N., Passi, V., Bryson, S., & Agras, W. S. (2005). Patterns of eating and abstinence in women treated

for bulimia nervosa. *International Journal of Eating Disorders*, *38*(4), 330–334. https://doi.org/10.1002/eat.20204

Shapiro, J. R., Berkman, N. D., Brownley, K. A., Sedway, J. A., Lohr, K. N., & Bulik, C. M. (2007). Bulimia nervosa treatment: A systematic review of randomized controlled trials. *International Journal of Eating Disorders*, *40*(4), 321–336. https://doi.org/10.1002/eat.20372

Shea, M. T., Elkin, I., Imber, S. D., Sotsky, S. M., Watkins, J. T., Collins, J. F., … Parloff, M. B. (1992). Course of depressive symptoms over follow-up: Findings from the National Institute of Mental Health Treatment of Depression Collaborative Research Program. *Archives of General Psychiatry*, *49*, 782–787.

Shrout, P. E., & Bolger, N. (2002). Mediation in experimental and nonexperimental studies: New procedures and recommendations. *Psychological Methods*, *7*(4), 422–445. https://doi.org/10.1037//1082-989X.7.4.422

Singer, J. D., & Willett, J. B. (2003). Treating TIME more flexibly. In *Applied Longitudinal Data Analysis* (2nd ed., pp. 138–181). Oxford: Oxford University Press. https://doi.org/10.1093/acprof:oso/9780195152968.003.0005

Sobel, M. E. (1982). Asymptotic confidence intervals for indirect effects in structural equation models. In S. Leinhart (Ed.), *Soociological Methodology* (pp. 290–312). San Francisco: Jossey-Bass.

Spangler, D. L., Baldwin, S. A., & Agras, W. S. (2004). An examination of the mechanisms of action in cognitive behavioral therapy for bulimia nervosa. *Behavior Therapy*, *35*(3), 537–560. Retrieved from http://www.sciencedirect.com/science/article/pii/S0005789404800315

StataCorp. (2013). Stata Statistical Software: Release 13. College Station, TX: StataCorp LP.

Steinhausen, H.-C., & Weber, S. (2009). The outcome of bulimia nervosa: Findings from one-quarter century of research. *American Journal of Psychiatry, 166*(12), 1331–1341. https://doi.org/10.1176/appi.ajp.2009.09040582

Sullivan, P. F., Bulik, C. M., Fear, J. L., & Pickering, A. (1998). Outcome of anorexia nervosa: A case-control study. *American Journal of Psychiatry*, *155*(7), 939–946.

Tabachnick, B. G., & Fidell, L. S. (2005). Multilevel linear modelling. In *Using Multivariate Statistics* (5th ed.). Essex, UK: Pearson.

Taubman, S. L., Robins, J. M., Mittleman, M. A., & Hernán, M. A. (2009). Intervening on risk factors for coronary heart disease: an application of the parametric g-formula. *International Journal of Epidemiology*, *38*(6), 1599–1611. https://doi.org/10.1093/ije/dyp192

Thackwray, D. E., Smith, M. C., Bodfish, J. W., & Meyers, A. W. (1993). A comparison of behavioral and cognitive-behavioral interventions for bulimia nervosa. *Journal of Consulting & Clinical Psychology*, *61*(4), 639–645.

Thompson-Brenner, H., Franko, D. L., Thompson, D. R., Grilo, C. M., Boisseau, C. L., Roehrig, J. P., … Wilson, G. T. (2013). Race/ethnicity, education, and treatment parameters as moderators and predictors of outcome in binge eating disorder. *Journal of Consulting and Clinical Psychology*, *81*(4), 710–721. https://doi.org/10.1037/a0032946

Thompson-Brenner, H., Shingleton, R. M., Thompson, D. R., Satir, D. A., Richards, L. K., Pratt, E. M., & Barlow, D. H. (2016). Focused vs. broad enhanced cognitive behavioral therapy for bulimia nervosa with comorbid borderline personality: A randomized controlled trial. *International Journal of Eating Disorders*, *49*(1), 36–49. https://doi.org/10.1002/eat.22468

Turner, H., Marshall, E., Stopa, L., & Waller, G. (2015). Cognitive-behavioural therapy for outpatients with eating disorders: Effectiveness for a transdiagnostic group in a routine clinical setting. *Behaviour Research and Therapy*, *68*, 70–75. https://doi.org/10.1016/j.brat.2015.03.001

Vall, E., & Wade, T. D. (2015). Predictors of treatment outcome in individuals with eating disorders: A systematic review and meta-analysis. *International Journal of Eating Disorders*, *48*(7), 946–971. https://doi.org/10.1002/eat.22411

von Soest, T., & Hagtvet, K. A. (2011). Mediation analysis in a latent growth curve modeling framework. *Structural Equation Modeling: A Multidisciplinary Journal*, *18*(2), 289–314. https://doi.org/10.1080/10705511.2011.557344

Waller, G., Corstorphine, E., Cordery, H., Hinrichsen, H., Lawson, R., Mountford, V., & Russell, K. (2007a). *Cognitive-behavioral therapy for the eating disorders: A comprehensive treatment guide.* Cambridge, UK: Cambridge University Press.

Waller, G., Corstorphine, E., Cordery, H., Hinrichsen, H., Lawson, R., Mountford, V., & Russell, K. (2007b). *Cognitive-Behavioral Therapy for the Eating Disorders: A Comprehensive Treatment Guide.*

Cambridge, UK: Cambridge University Press.
Waller, G., Evans, J., & Pugh, M. (2013). Food for thought: A pilot study of the pros and cons of changing eating patterns within cognitive-behavioural therapy for the eating disorders. *Behaviour Research and Therapy*, *51*(9), 519–525. https://doi.org/10.1016/j.brat.2013.06.001
Waller, G., & Mountford, V. A. (2015). Weighing patients within cognitive-behavioural therapy for eating disorders: How, when and why. *Behaviour Research and Therapy*, *70*, 1–10. https://doi.org/10.1016/j.brat.2015.04.004
Weissman, M. M. (2006). A brief history of interpersonal psychotherapy. *Psychiatric Annals*, *36*(8), 553–557.
Weissman, M. M. (2007). Cognitive therapy and interpersonal psychotherapy: 30 years later. *American Journal of Psychiatry*, *164*(5), 693–696. https://doi.org/10.1176/appi.ajp.164.5.693
Wilfley, D. E., Welch, R., Stein, R. I., Spurrell, E. B., Cohen, L. R., Saelens, B. E., … Matt, G. E. (2002). A randomized comparison of group cognitive-behavioral therapy and group interpersonal psychotherapy for the treatment of overweight individuals with binge-eating disorder. *Archives of General Psychiatry*, *59*(8), 713–721. https://doi.org/10.1001/archpsyc.59.8.713
Willett, J. B., & Sayer, A. G. (1994). Using covariance structure analysis to detect correlates and predictors of individual change over time. *Psychological Bulletin*, *116*(2), 363–381. https://doi.org/10.1037/0033-2909.116.2.363
Wilson, G. T., Fairburn, C. G., Agras, W. S., Walsh, B. T., & Kraemer, H. (2002). Cognitive-behavioral therapy for bulimia nervosa: Time course and mechanisms of change. *Journal of Consulting and Clinical Psychology*, *70*(2), 267–274. https://doi.org/10.1037/0022-006X.70.2.267
Wilson, G. T., Rossiter, E., Kleifield, E. I., & Lindholm, L. (1986). Cognitive-behavioral treatment of bulimia nervosa: A controlled evaluation. *Behaviour Research and Therapy*, *24*(3), 277–288. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/3460591
Wonderlich, S. A., Peterson, C. B., Crosby, R. D., Smith, T. L., Klein, M. H., Mitchell, J. E., & Crow, S. J. (2014). A randomized controlled comparison of integrative cognitive-affective therapy (ICAT) and enhanced cognitive-behavioral therapy (CBT-E) for bulimia nervosa. *Psychological Medicine*, *44*(3), 543–553. https://doi.org/10.1017/S0033291713001098
Wu, W., & Jia, F. (2013). A new procedure to test mediation with missing data through nonparametric bootstrapping and multiple imputation. *Multivariate Behavioral Research*, *48*(5), 663–691. https://doi.org/10.1080/00273171.2013.816235
Zendegui, E. A., West, J. A., & Zandberg, L. J. (2014). Binge eating frequency and regular eating adherence: The role of eating pattern in cognitive behavioral guided self-help. *Eating Behaviors*, *15*(2), 241–243. https://doi.org/10.1016/j.eatbeh.2014.03.002
Zhang, Z., Zyphur, M. J., & Preacher, K. J. (2009). Testing multilevel mediation using hierarchical linear models: Problems and solutions. *Organizational Research Methods*, *12*(4), 695–719. https://doi.org/10.5465/AMBPP.2008.33716518
Zipfel, S., Wild, B., Groß, G., Friederich, H.-C., Teufel, M., Schellberg, D., … Herzog, W. (2014). Focal psychodynamic therapy, cognitive behaviour therapy, and optimised treatment as usual in outpatients with anorexia nervosa (ANTOP study): Randomised controlled trial. *The Lancet*, *383*(9912), 127–137. https://doi.org/10.1016/S0140-6736(13)61746-8
Zunker, C., Peterson, C. B., Crosby, R. D., Cao, L., Engel, S. G., Mitchell, J. E., & Wonderlich, S. A. (2011). Ecological momentary assessment of bulimia nervosa: Does dietary restriction predict binge eating? *Behaviour Research and Therapy*, *49*(10), 714–717. https://doi.org/10.1016/j.brat.2011.06.006

Relevant Background Papers Resulting from Main Efficacy Trial

Appendix 1: Main Efficacy Trial Paper (Fairburn et al., 2015)

Behaviour Research and Therapy 70 (2015) 64–71

Contents lists available at ScienceDirect

Behaviour Research and Therapy

journal homepage: www.elsevier.com/locate/brat

A transdiagnostic comparison of enhanced cognitive behaviour therapy (CBT-E) and interpersonal psychotherapy in the treatment of eating disorders

Christopher G. Fairburn [a, *], Suzanne Bailey-Straebler [a], Shawnee Basden [a], Helen A. Doll [b], Rebecca Jones [c], Rebecca Murphy [a], Marianne E. O'Connor [a], Zafra Cooper [a]

[a] Oxford University, Department of Psychiatry, UK
[b] Department of Population Health and Primary Care, Norwich Medical School, University of East Anglia, UK
[c] London School of Hygiene and Tropical Medicine, UK

ARTICLE INFO

Article history:
Received 20 February 2015
Received in revised form
1 April 2015
Accepted 20 April 2015
Available online 22 April 2015

Keywords:
Eating disorders
Treatment
Cognitive behaviour therapy
Interpersonal psychotherapy

ABSTRACT

Eating disorders may be viewed from a transdiagnostic perspective and there is evidence supporting a transdiagnostic form of cognitive behaviour therapy (CBT-E). The aim of the present study was to compare CBT-E with interpersonal psychotherapy (IPT), a leading alternative treatment for adults with an eating disorder. One hundred and thirty patients with any form of eating disorder (body mass index >17.5 to <40.0) were randomized to either CBT-E or IPT. Both treatments involved 20 sessions over 20 weeks followed by a 60-week closed follow-up period. Outcome was measured by independent blinded assessors. Twenty-nine participants (22.3%) did not complete treatment or were withdrawn. At post-treatment 65.5% of the CBT-E participants met criteria for remission compared with 33.3% of the IPT participants ($p < 0.001$). Over follow-up the proportion of participants meeting criteria for remission increased, particularly in the IPT condition, but the CBT-E remission rate remained higher (CBT-E 69.4%, IPT 49.0%; $p = 0.028$). The response to CBT-E was very similar to that observed in an earlier study. The findings indicate that CBT-E is potent treatment for the majority of outpatients with an eating disorder. IPT remains an alternative to CBT-E, but the response is less pronounced and slower to be expressed.
Current controlled trials: ISRCTN 15562271.

© 2015 Published by Elsevier Ltd.

1. Introduction

Most studies of mental disorders have focused on specific disorders in isolation. This strategy has been criticised as it complicates the identification of psychopathological processes that operate across disorders (Insel, 2014; Insel et al., 2010) and it makes it difficult to identify treatments that are potentially "transdiagnostic" in their clinical range.

In 2003 we suggested that there would be value in viewing eating disorders from a "transdiagnostic" perspective (Fairburn, Cooper, & Shafran, 2003). We proposed that eating disorder psychopathology is maintained by a largely common set of mechanisms, and that treatments capable of addressing these mechanisms should be effective across the various eating disorder presentations. Accordingly, we modified the leading evidence-based treatment for the eating disorders, a cognitive behavioural treatment for bulimia nervosa (CBT-BN) (Fairburn, 1981; Fairburn, Marcus, & Wilson, 1993), to make it suitable for all forms of eating disorder while at the same time we attempted to make it more potent (Cooper & Fairburn, 2011; Fairburn et al., 2003). The resulting treatment, enhanced cognitive behaviour therapy or CBT-E, has been investigated as a treatment for anorexia nervosa (Dalle Grave, Calugi, Conti, Doll, & Fairburn, 2013; Dalle Grave, Calugi, Doll, & Fairburn, 2013; Dalle Grave, Calugi, Ghoch, Conti M, & Fairburn, 2014; Fairburn et al., 2013; Zipfel et al., 2014) and bulimia nervosa (Poulsen et al., 2014; Wonderlich et al., 2014), and in two transdiagnostic samples (Fairburn et al., 2009) that have included cases of bulimia nervosa, binge eating disorder and the various unspecified eating disorder presentations seen in adults (collectively termed "eating disorder not otherwise specified" in

* Corresponding author. Oxford University, Department of Psychiatry, Warneford Hospital, Oxford OX3 7JX, UK.
E-mail address: credo@medsci.ox.ac.uk (C.G. Fairburn).
URL: http://www.credo-oxford.com

http://dx.doi.org/10.1016/j.brat.2015.04.010
0005-7967/© 2015 Published by Elsevier Ltd.

DSM-IV (American Psychiatric Association, 1994)). The findings indicate that CBT-E is indeed transdiagnostic in its clinical range, but its relative and absolute effectiveness remain to be established. Neither are clear because there have been differences across the studies in the way that CBT-E has been implemented.

The primary aim of the present study was to compare the effects of CBT-E with those of interpersonal psychotherapy (IPT), the leading alternative to cognitive behaviour therapy as a treatment for adult outpatients with an eating disorder. There is evidence supporting the use of IPT in bulimia nervosa (Agras, Walsh, Fairburn, Wilson, & Kraemer, 2000; Fairburn, Jones, Peveler, Hope, & O'Connor, 1993), binge eating disorder (Wilfley et al., 1993, 2002; Wilson, Wilfley, Agras, & Bryson, 2010) and, to a lesser extent, anorexia nervosa (Carter et al., 2011; McIntosh et al., 2005). IPT has not previously been tested in a broad transdiagnostic patient sample nor has it been compared with CBT-E. The second aim was to determine whether the findings of the original and largest study of CBT-E (Fairburn et al., 2009) could be replicated when an equivalent inclusive patient sample was recruited and CBT-E was implemented in the same way.

2. Method

2.1. Design

A randomized controlled trial was conducted at a catchment area-based eating disorder clinic in the UK. Eligible patients were randomized either to CBT-E or IPT delivered over 20 weeks. They were assessed before and after treatment and then entered a closed follow-up period during which they were assessed at 20, 40 and 60 weeks post-treatment. During the follow-up period they received no further treatment unless it was judged necessary on clinical grounds. The study was restricted to eating disorder patients whose body mass index (BMI) was over 17.5 and under 40.0. The study was approved by the local human subjects committee.

2.2. Recruitment

The sample was recruited from consecutive referrals from family doctors and other clinicians to a long-established eating disorder clinic serving central Oxfordshire in the UK. Fig. 1 shows the CONSORT diagram.

Recruitment was designed to be "inclusive" with few exclusion criteria. The eligibility criteria were having an eating disorder that required treatment, as judged both by the referring clinician and, subsequently, by a senior eating disorder specialist (ZC or CGF); being aged 18–65 years; having a BMI over 17.5 and under 40.0; and providing written informed consent after receiving a detailed description of the study. The exclusion criteria were prior receipt of a treatment closely resembling CBT-E or IPT (N = 13); a co-existing general psychiatric disorder that precluded eating disorder-focused treatment (N = 16); medical instability or pregnancy (N = 13); and not being available for treatment (N = 13). Examples of co-existing general psychiatric disorders that precluded immediate entry into the trial were severe clinical depression, bipolar disorder and marked agoraphobia. Participants who were receiving ongoing psychiatric treatment were weaned off this prior to entering the study, the exception being clinically warranted antidepressant medication (N = 57) which was kept stable during treatment.

2.3. Treatment

Both treatments were delivered in 20 50-min sessions, preceded by one 90-min preparatory session, and followed by a review session 20 weeks after the completion of treatment.

Enhanced cognitive behaviour therapy – CBT-E is a psychological treatment designed to address eating disorder psychopathology whatever the eating disorder diagnosis. It is personalised to match the eating disorder psychopathology of the individual patient (Fairburn, 2008). The default ("focused") form of the treatment was used. It concentrates exclusively on the modification of eating habits, weight-control behaviour, and concerns about eating, shape and weight.

Interpersonal psychotherapy – IPT is a short-term psychological treatment designed to identify and address current interpersonal problems (Klerman, Weissman, Rounsaville, & Chevron, 1984). It was originally developed as a treatment for depression but in the 1980s it was adapted by the first author to make it suitable for the treatment of bulimia nervosa (Fairburn, 1993). This form of IPT has been shown to be as effective as CBT-BN but it is substantially slower to achieve its results (Agras et al., 2000; Fairburn, Marcus, et al., 1993; Fairburn, Jones, et al., 1993). IPT has also been successfully used in the treatment of binge eating disorder (Wilfley et al., 2002; Wilson et al., 2010).

The form of IPT used in the present study has been described in detail elsewhere (Fairburn, 1993; Murphy, Straebler, Basden, Cooper, & Fairburn, 2012). It closely resembled IPT for depression but with minor adjustments so that it could be applied to people with any eating disorder (Murphy et al., 2012). It included all IPT's key strategies and procedures including role play.

Therapists – Three therapists took part, two were clinical psychologists and one was a psychiatric nurse practitioner. All had generic clinical experience and experience treating patients with eating disorders. One had prior experience implementing CBT-E and another had experience using IPT. Each therapist received six months' training from CGF and ZC in both CBT-E and IPT prior to the formal start of the study. During the study weekly supervision meetings were led by ZC and CGF with close and equal attention being paid to the quality of implementation of the two treatments. All the treatment sessions were recorded and each week selected sessions were audited by the supervisors.

Treatment fidelity – The quality of delivery of the two treatments was assessed by independent raters using an adaptation of an instrument developed for a prior comparison of CBT-BN and IPT (Loeb et al., 2005). There were 20 CBT-E items and these addressed general features of CBT-E, stage-specific features and the use of non-cognitive behavioural techniques. The 20 IPT items were equivalent in their focus but IPT-relevant. All the items were rated using the same seven-point rating scheme. Following completion of the treatment phase of the study 120 treatment sessions were selected at random from the early, middle, and late stages of both treatments and rated blind by four postgraduate research assistants who had received training in the use of the rating scale.

2.4. Assessment

Eating disorder features – These were assessed using the 16th edition of the Eating Disorder Examination interview (EDE) (Fairburn, Cooper, & O'Connor, 2008) and its self-report version (EDE-Q6.0) (Fairburn & Beglin, 2008). The EDE was administered by assessors who were trained and supervised by MO'C, an expert on the instrument. The assessors were kept blind to the participants' treatment condition and had no involvement with treatment. Operational DSM-IV diagnoses of bulimia nervosa and binge eating disorder (American Psychiatric Association, 1994) were generated from the EDE ratings (see Supplementary Table 1), the remaining participants being given the diagnosis "other eating disorder" (OED). Two outcome variables were created from the EDE ratings; the severity of eating disorder features as measured by the global EDE score (continuous), and the primary outcome variable of being

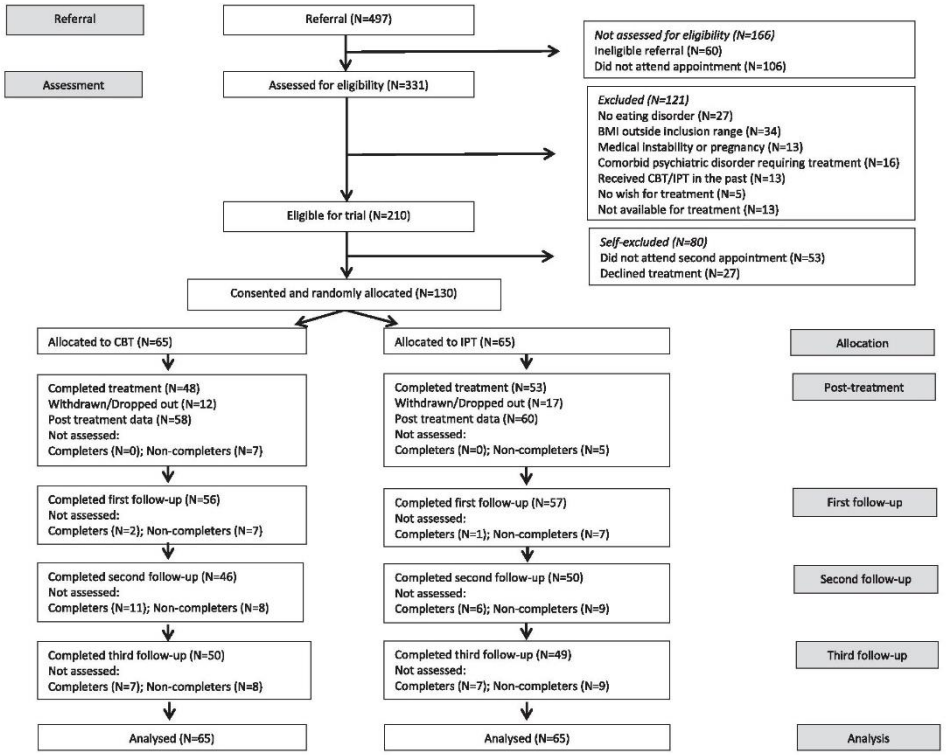


Fig. 1. The CONSORT diagram.

in "remission" defined as a global EDE score less than 1 standard deviation (SD) above the community mean (categorical); i.e., below 1.74 (Beglin, 1990). Normative comparisons of this type are widely used to identify clinically significant change (Kazdin, 2003; Kendall, Marrs-Garcia, Nath, & Sheldrick, 1999).

Psychosocial impairment secondary to eating disorder features – This was measured using the Clinical Impairment Assessment questionnaire (CIA) (Bohn & Fairburn, 2008; Bohn et al., 2008).

General psychiatric features – The Structured Clinical Interview for DSM-IV (First, Gibbon, & Williams, 1997) was used at baseline to identify the presence of co-existing general psychiatric disorders, and the level of depressive features was measured using the Beck Depression Inventory (BDI) (Beck, Ward, Mendelson, Mock, & Erbaugh, 1961).

Treatment suitability and expectancy – These were assessed after two treatment sessions and suitability was assessed again at the end of treatment. Both were measured using widely employed patient-rated visual analogue scales (Agras et al., 2000).

2.5. Power and sample size

Sample size calculations were performed a priori with respect to the comparison of CBT-E and IPT, and they were on an intent-to-treat (ITT) basis. It was calculated that a sample size of 65 participants per treatment condition was required to provide 80% power at two-sided $p < 0.05$ to detect a difference at the end of treatment in global EDE change of 0.45 points, assuming a SD of global EDE change scores of 1.0 (Agras et al., 2000) (i.e., a moderate effect size of 0.5) and also to detect a difference between the two conditions of at least 25% in the categorical outcome measure.

2.6. Randomization

Participants were allocated to the two treatment conditions by HAD (who had no involvement in recruitment) using a computer-based minimization algorithm to balance on gender, eating disorder diagnosis, BMI, and the need for participants to remain on psychotropic medication. When groups were evenly balanced, pre-prepared blocked randomization lists of varying size were used to allocate participants to the two conditions.

2.7. Statistical methods

The primary analysis was an ITT analysis which included all randomised participants. Each continuous measure was analysed using a separate linear mixed model with the scores from each time point treated as a repeated measures outcome. Fixed effects of treatment (IPT vs CBT-E), time (baseline, end of treatment, and 20, 40 and 60 weeks post-treatment), and the interaction between treatment and time were specified, and the estimated baseline score was constrained to be identical in the two treatment conditions. This is equivalent to adjusting for baseline and permitting the

relationship between baseline and follow up scores to differ at each time point, but offers the additional advantage that the data from all participants contribute to the analysis, even if there are missing data at follow up. An unstructured residual covariance matrix was used to allow for correlations within participants between the repeated measures over time. Each binary outcome was analysed in a similar fashion using a mixed effects logistic regression model with the binary indicator from each time point as the repeated measures outcome. Fixed effects were as for the continuous outcomes. A random effect of patient was specified to allow for correlations between repeated measures within participants. Statistical significance was taken at the 5% level ($p < 0.05$).

A per protocol analysis was also conducted using the same data analytic strategy as the ITT analysis. It was restricted to those participants who were considered by their therapist to have fully completed treatment.

All the analyses were undertaken by RJ using Stata v13 (StataCorp, College Station, Texas).

3. Results

3.1. Sample

One hundred and thirty eligible participants were recruited and randomized between 2006 and 2011. Their diagnoses were as follows: bulimia nervosa – 53 participants (40.8%); binge eating disorder – 8 participants (6.2%); and "other eating disorder" – 69 participants (53.1%). Sixty-five were randomized to CBT-E and 65 to IPT. The characteristics of the sample are shown in Table 1.

3.2. Suitability, expectancy, therapy quality and attrition

The ratings of treatment expectancy were high and did not differ between the two treatments (expectancy, mean and SD – IPT 66.1, 18.6; CBT-E 68.1, 20.5). The same was true of treatment suitability (beginning, mean and SD – IPT 69.2, 18.7; CBT-E 76.6, 19.7; end, mean and SD – IPT 75.5, 21.3; CBT-E 78.2, 24.4). The ratings of treatment fidelity were also high with over two-thirds of the sessions in both treatments being rated as "excellent" (i.e., ratings of 6 or 7 on the 1 to 7 point rating scale).

Twenty-nine participants (22.3%) did not complete all 20 sessions of treatment or were withdrawn because of lack of response. The non-completion figures by diagnosis were as follows: bulimia nervosa – 32.1% (17/53); binge eating disorder – 0% (0/8); other eating disorder – 17.4% (12/69).

3.3. Relative effects of CBT-E and IPT

3.3.1. At post-treatment

At post-treatment the levels of eating disorder and general psychopathology had decreased in both treatment conditions but the changes were significantly greater among the CBT-E participants (see Table 2 and Fig. 2). The observed proportion of CBT-E participants in remission (i.e., a global EDE score below 1.74) was

Table 1
Characteristics of the sample at baseline.

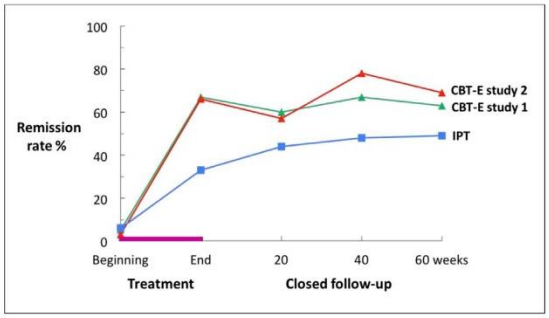
	IPT (N = 65)		CBT-E (N = 65)		All Patients (N = 130)	
	Mean	**(SD)**	**Mean**	**(SD)**	**Mean**	**(SD)**
Age (years)	26.8	(8.8)	24.9	(6.4)	25.9	(7.7)
	N	**(%)**	**N**	**(%)**	**N**	**(%)**
Female	63	(96.9)	64	(98.5)	127	(97.7)
Ethnicity:						
White	60	(92.3)	64	(98.5)	124	(95.4)
Black British	1	(1.5)	0	(0.0)	1	(0.8)
Asian Chinese	0	(0.0)	1	(1.5)	1	(0.8)
Asian British	1	(1.5)	0	(0.0)	1	(0.8)
Mixed	3	(4.6)	0	(0.0)	3	(2.3)
Marital Status:						
Single, never married	52	(80.0)	58	(89.2)	110	(84.6)
Married or living as such	12	(18.5)	6	(9.2)	18	(13.9)
Separated or divorced	1	(1.5)	1	(1.5)	2	(1.5)
Occupational Social Class:						
Higher	18	(27.7)	8	(12.3)	26	(20.0)
Intermediate	5	(7.7)	9	(13.9)	14	(10.8)
Lower	10	(15.4)	11	(16.9)	21	(16.2)
Unclassifiable	2	(3.1)	2	(3.1)	4	(3.1)
Student	30	(46.2)	35	(53.9)	65	(50.0)
DSM-IV Eating Disorder Status:						
Bulimia nervosa	28	(43.1)	25	(38.5)	53	(40.8)
Binge eating disorder	4	(6.2)	4	(6.2)	8	(6.2)
Other eating disorder	33	(50.8)	36	(55.4)	69	(53.1)
History of anorexia nervosa	15	(23.8)	21	(32.8)	36	(28.4)
	Mean	**(SD)**	**Mean**	**(SD)**	**Mean**	**(SD)**
Duration of eating disorder (years)	11.4	(9.6)	8.4	(7.3)	9.9	(8.6)
Lowest adult BMI (kg/m^2)	19.3	(2.9)	18.9	(2.8)	19.1	(2.9)
Highest adult BMI (kg/m^2)	26.6	(5.6)	26.2	(5.4)	26.4	(5.5)
	N	**(%)**	**N**	**(%)**	**N**	**(%)**
Current Comorbid Diagnoses (SCID):						
Major depressive episode	6	(9.2)	9	(13.9)	15	(11.5)
Any anxiety disorder	15	(23.1)	14	(21.5)	29	(22.3)
Substance abuse	7	(10.8)	2	(3.1)	9	(6.9)
Any Axis 1 disorder	25	(38.5)	22	(34.4)	47	(36.4)
Need for psychotropic medication	27	(41.5)	30	(46.2)	57	(43.9)

BMI – Body mass index; SCID – Structured Clinical Interview for DSM-IV Axis I Disorders (First et al., 1997).
Missing values – There were 2 missing values for history of anorexia nervosa (IPT: 1 missing; CBT: 1 missing) and 1 missing value for any axis I disorder (IPT: 0 missing; CBT: 1 missing). The denominator for percentages is the number of non-missing values.

Table 2
Main clinical features at baseline, post-treatment and at 60-week post-treatment follow-up.

	Baseline				Post-treatment				60-week Follow-up			
	IPT (N = 65)		CBT (N = 65)		IPT (N = 60)		CBT (N = 58)		IPT (N = 49)		CBT (N = 49)	
	Mean	**(SD)**	**Mean**	**(SD)**	**Mean**	**(SD)**	**Mean**	**(SD)**	**Mean**	**(SD)**	**Mean**	**(SD)**
Body mass index (kg/m^2)	22.8	(4.2)	22.9	(4.4)	23.6	(4.5)	23.5	(4.1)	24.8	(5.1)	24.1	(4.8)
Eating disorder psychopathology (EDE)												
Global score	3.52	(1.05)	3.59	(1.01)	2.37	(1.25)	1.57	(1.25)	1.83	(1.28)	1.51	(1.20)
Dietary restraint	3.70	(1.32)	3.71	(1.09)	2.46	(1.61)	1.08	(1.35)	1.71	(1.53)	1.32	(1.53)
Eating concern	2.86	(1.11)	2.81	(1.27)	1.83	(1.40)	1.12	(1.21)	1.24	(1.23)	0.97	(1.09)
Shape concern	4.03	(1.32)	4.08	(1.38)	2.53	(1.34)	1.99	(1.43)	2.30	(1.57)	1.98	(1.48)
Weight concern	3.47	(1.51)	3.77	(1.31)	2.67	(1.65)	2.09	(1.67)	2.08	(1.60)	1.78	(1.41)
Other features												
Secondary impairment (CIA)	30.0	(8.3)	30.5	(8.6)	19.6	(12.4)	13.9	(10.4)	12.6	(10.8)	12.4	(12.1)
Depressive features (BDI)	22.8	(10.9)	21.2	(10.8)	14.0	(12.3)	11.8	(11.0)	11.8	(12.6)	12.7	(11.8)
	N	**(%)**	**N**	**(%)**	**N**	**(%)**	**N**	**(%)**	**N**	**(%)**	**N**	**(%)**
EDE global score < 1 SD above the community mean (<1.74)	4	(6.2)	2	(3.1)	20	(33.3)	38	(65.5)	24	(49.0)	34	(69.4)
Eating disorder behaviour (EDE)												
Objective bulimic episodes ≥1	51	(78.5)	54	(83.1)	38	(63.3)	25	(43.1)	20	(40.8)	19	(38.8)
Self-induced vomiting ≥1	42	(64.6)	41	(63.1)	31	(51.7)	22	(37.9)	19	(38.8)	19	(38.8)
Laxative-taking ≥1	17	(26.2)	7	(10.8)	12	(20.0)	1	(1.7)	5	(10.2)	1	(2.0)
Absence of all of the above	8	(12.3)	6	(9.2)	13	(21.7)	26	(44.8)	23	(46.9)	22	(44.9)
Cessation of binge eating and purging if present at baseline (N, %)	–	–	–	–	7/52	(13.5)	22/53	(41.5)	16/41	(39.0)	18/45	(40.0)
	Median	**(IQR)**	**Median**	**(IQR)**	**Median**	**(IQR)**	**Median**	**(IQR)**	**Median**	**(IQR)**	**Median**	**(IQR)**
Objective bulimic episodes (N)	14	(3, 28)	11	(4, 28)	7	(0, 18)	0	(0, 5)	0	(0, 4)	0	(0, 2)
Self-induced vomiting (N)	14	(0, 40)	5	(0, 25)	1.5	(0, 19)	0	(0, 5)	0	(0, 7)	0	(0, 3)
Laxative-taking (N)	0	(0, 1)	0	(0, 0)	0	(0, 0)	0	(0, 0)	0	(0, 0)	0	(0, 0)

EDE – Eating Disorder Examination (Fairburn et al., 2008); CIA – Clinical Impairment Assessment (Bohn & Fairburn, 2008; Bohn et al., 2008); BDI – Beck Depression Inventory (Beck et al., 1961); BMI – Body mass index.
Missing values – Immediately post-treatment there were a further 4 missing values for BMI, CIA and BDI (IPT: 1; CBT: 3). At 60-week post-treatment, there were a further 6 missing values for BMI (IPT: 4; CBT: 2), 14 missing for CIA (IPT: 7; CBT: 7) and 10 missing for BDI (IPT: 4; CBT: 6).



CBT-E study 1 - Focused CBT-E. (Data from the original transdiagnostic study (Fairburn et al., 2009) analysed using the statistical approach employed in the present study).

CBT-E study 2 - Focused CBT-E. (Data from present study).

IPT - Interpersonal psychotherapy. (Data from present study).

Fig. 2. Intent-to-treat remission rates in the present study and the earlier one (Fairburn et al., 2009).

almost twice that of those who received IPT (CBT-E – 38/58, 65.5%; IPT – 20/60, 33.3%; adjusted OR 8.8; 95% CI 2.6 to 29.5; $p < 0.001$; see Table 3). Almost half of the CBT-E participants (44.8%, 26/58) reported no binge eating, vomiting or laxative misuse at the end of treatment compared with 21.7% (13/60) in the IPT condition (adjusted OR 6.7; 95% CI 1.9 to 23.6; $p = 0.003$).

The changes observed were greater among the participants who completed treatment but the relative effects of the two treatments were similar (see Supplementary Tables 1 and 2). Three-quarters of those who completed CBT-E were in remission compared with just over a third of those who completed IPT (CBT-E – 36/48, 75.0%; IPT – 20/53, 37.7%; adjusted OR 13.0; 95% CI 3.4 to 49.4; $p < 0.001$; see Table 3).

3.3.2. *Over the 60-week post-treatment follow-up*

There was high compliance with the 60-week period of post-treatment follow-up with 79.2% (309/390) of the assessments being successfully completed. Few of the participants required additional treatment. Of the 118 participants who were assessed at the end of treatment and then entered follow-up, seven had further treatment at some point in the 60 weeks and a further seven had up to five brief "booster" sessions.

The changes observed during treatment were well maintained in the CBT-E condition (see Table 2 and Fig. 2). For example, at the

Table 3
Intention-to-treat analysis at post-treatment and at 60-week post-treatment follow-up.

	Effect estimates for CBT-E vs IPT post-treatment			Effect estimate for CBT-E vs IPT 60-week follow-up		
	Difference	**(95% CI)**	**p value**	**Difference**	**(95% CI)**	**p value**
Body mass index (kg/m^2)	0.14	(−0.42–0.71)	0.621	−0.59	(−1.50–0.32)	0.204
Eating disorder psychopathology (EDE)						
Global score	−0.81	(−1.23–−0.40)	<0.001	−0.28	(−0.74–0.18)	0.230
Dietary restraint	−1.39	(−1.89–−0.89)	<0.001	−0.42	(−0.98–0.14)	0.142
Eating concern	−0.69	(−1.14–−0.23)	0.003	−0.23	(−0.68–0.21)	0.304
Shape concern	−0.54	(−0.99–−0.08)	0.021	−0.29	(−0.85–0.26)	0.303
Weight concern	−0.66	(−1.21–−0.11)	0.019	−0.28	(−0.82–0.26)	0.308
Other features						
Secondary Impairment (CIA)	−6.22	(−10.10–−2.33)	0.002	−1.33	(−5.72–3.05)	0.551
Depressive features (BDI)	−1.76	(−5.86–2.34)	0.399	1.12	(−3.40–5.63)	0.628
	Odds ratio	**(95% CI)**	**p value**	**Odds ratio**	**(95% CI)**	**p value**
EDE severity score < 1 SD above the community mean (<1.74)	8.75	(2.59–29.54)	<0.001	4.20	(1.17–15.14)	0.028
Eating disorder behaviour (EDE)						
Objective bulimic episodes ≥1	0.24	(0.08–0.75)	0.014	0.80	(0.24–2.66)	0.714
Self-induced vomiting ≥1	0.22	(0.04–1.09)	0.064	0.84	(0.15–4.79)	0.843
Laxative-taking ≥1	0.03	(0.00–0.54)	0.018	0.11	(0.00–2.82)	0.183
Absence of all of the above	6.68	(1.89–23.59)	0.003	1.04	(0.29–3.67)	0.954

EDE – Eating Disorder Examination (Fairburn et al., 2008); CIA – Clinical Impairment Assessment (Bohn & Fairburn, 2008; Bohn et al., 2008); BDI – Beck Depression Inventory (Beck et al., 1961).
Estimated treatment differences are from longitudinal mixed effects linear or logistic regression models which included data from all participants who were randomised (IPT: 65; CBT: 65), account for missing data and repeated measures on individuals over time and adjust for baseline values of the outcome in question.

end of treatment the mean global EDE score was 1.57 (SD = 1.25) and at 60-week follow-up it was 1.51 (SD = 1.20). Over the same period the proportion of participants in remission increased slightly from 65.5% (38/58) at the end of treatment to 69.4% (34/49) at 60-week follow-up (see Fig. 2).

In the IPT condition the level of psychopathology fell over the period of follow-up with the result that many of the post-treatment differences between CBT-E and IPT were no longer statistically significant at 60-week follow-up (e.g. adjusted mean difference in global EDE score −0.28, 95% CI -0.74 to 0.18, p = 0.23; see Table 3), an exception being the primary outcome variable of being in remission which remained higher among the CBT-E participants than among those who received IPT (CBT-E – 34/49, 69.4%; IPT – 24/49, 49.0%; adjusted OR: 4.2; 95% CI: 1.2 to 15.1; p = 0.028; see Table 3 and Fig. 2).

The remission rates during the period of follow-up were higher among those who completed treatment but the relative effects of the two treatments remained similar (see Supplementary Tables 2 and 3). The proportion of participants in remission at 60-week follow-up was 70.0% (28/40) in the CBT-E condition and 50.0% (23/46) among those who received IPT (adjusted OR 3.8; 95% CI 1.0 to 14.6; p = 0.049).

3.4. Absolute effects of CBT-E

Fig. 2 shows the proportion of participants at each assessment point with a global EDE-Q score below 1.74, together with the equivalent data for CBT-E (focused version) from the earlier study (Fairburn et al., 2009) analysed using the statistical approach described above. It can be seen that in both studies the effects of CBT-E were substantial and well maintained. Furthermore, the magnitude of the response to CBT-E was remarkably similar in the two studies; the remission rates at the end of treatment being 67% in study 1 and 66% in study 2, and at 60-week post-treatment follow-up 63% and 69% respectively.

4. Discussion

This is the second randomised controlled trial to focus on the effects of CBT-E in a broad transdiagnostic sample of adult outpatients with an eating disorder (BMI >17.5 to <40.0). The first study recruited patients from two well-established eating disorder clinics in England, one in Leicestershire and the other in Oxfordshire (Fairburn et al., 2009). Both receive the great majority of referrals from the surrounding rural and urban areas. The present study recruited from the Oxfordshire-based clinic alone. In both studies care was taken to ensure that the great majority of referrals with a sufficiently severe eating disorder were eligible to take part. The resulting participant groups were very similar to each other. Both included the majority of those who were suitable and willing to be treated as outpatients, and they comprised patients with bulimia nervosa, binge eating disorder, the many and varied unspecified eating disorder presentations (Fairburn et al., 2007), and less underweight cases of anorexia nervosa (BMI > 17.5).

The first study compared the focused and broad forms of CBT-E, and it included a delayed treatment comparison condition (Fairburn et al., 2009). The present study involved a comparison of the focused form of CBT-E (the default version) with IPT, the leading alternative psychological treatment for adults with an eating disorder. Great care was taken to avoid any possible allegiance effects although none was likely as both treatments were devised by our group and we are known to be advocates of both CBT-E and IPT for eating disorders, The therapists received equivalent amounts of training in the two treatments and equivalent emphasis was placed upon them in the weekly supervision meetings. The independent ratings of treatment fidelity indicated that both were well implemented and there was no difference between the treatments in the ratings of suitability and expectancy.

The main aim of the study was to determine the relative effects of CBT-E and IPT. It was found that two-thirds of those who received CBT-E were in remission at the end of treatment compared with one-third of those who received IPT (CBT-E 66%; IPT 33%). This difference between the two treatments remained statistically significant 60 weeks later despite significant additional improvement in the IPT condition (CBT-E 69%; IPT 49%). The different temporal pattern of response to the two treatments (shown in Fig. 2) resembles that seen in the earlier studies of bulimia nervosa (Agras et al., 2000; Fairburn, Marcus, et al., 1993; Fairburn, Jones, et al., 1993). It is perhaps not surprising that CBT-E operates more rapidly than IPT given that it directly addresses the eating disorder

Eating Disorders, 44, 647–654.
Cooper, Z., & Fairburn, C. G. (2011). The evolution of "enhanced" cognitive behavior therapy for eating disorders: learning from treatment nonresponse. *Cognitive and Behavioral Practice, 18*, 394–402.
Dalle Grave, R., Calugi, S., Conti, M., Doll, H., & Fairburn, C. G. (2013). Inpatient cognitive behaviour therapy for anorexia nervosa: a randomized controlled trial. *Psychotherapy and Psychosomatics, 82*, 390–398.
Dalle Grave, R., Calugi, S., Doll, H. A., & Fairburn, C. G. (2013). Enhanced cognitive behaviour therapy for adolescents with anorexia nervosa: an alternative to family therapy? *Behaviour Research and Therapy, 51*, R9–R12.
Dalle Grave, R., Calugi, S., Ghoch, M. E., Conti, M., & Fairburn, C. G. (2014). Inpatient cognitive behavior therapy for adolescents with anorexia nervosa: immediate and longer-term effects. *Frontiers in Psychiatry, 5*, 1–7.
Fairburn, C. G. (1981). A cognitive behavioural approach to the treatment of bulimia. *Psychological Medicine, 11*, 707–711.
Fairburn, C. G. (1993). Interpersonal psychotherapy for bulimia nervosa. In G. L. Klerman, & M. M. Weissman (Eds.), *New applications of interpersonal psychotherapy* (pp. 353–378). Washington DC: American Psychiatric Press.
Fairburn, C. G. (2008). *Cognitive behavior therapy and eating disorders*. New York: Guilford Press.
Fairburn, C., & Beglin, S. J. (2008). Eating disorder examination questionnaire (EDE-Q 6.0). In C. Fairburn (Ed.), *Cognitive behavior therapy and eating disorders* (pp. 309–313). New York: Guilford Press.
Fairburn, C. G., Cooper, Z., Bohn, K., O'Connor, M. E., Doll, H. A., & Palmer, R. L. (2007). The severity and status of eating disorder NOS: Implications for DSM-V. *Behaviour Research and Therapy, 45*, 1705–1715.
Fairburn, C. G., Cooper, Z., Doll, H., O'Connor, M., Bohn, K., Hawker, D., et al. (2009). Transdiagnostic cognitive-behavioral therapy for patients with eating disorders: a two-site trial with 60-week follow-up. *American Journal of Psychiatry, 166*, 311–319.
Fairburn, C. G., Cooper, Z., Doll, H. A., O'Connor, M. E., Palmer, R. L., & Dalle Grave, R. (2013). Enhanced cognitive behaviour therapy for adults with anorexia nervosa: a UK-Italy study. *Behaviour Research and Therapy, 51*, R2–R8.
Fairburn, C. G., Cooper, Z., & O'Connor, M. (2008). Eating disorder examination (Edition 16.0D). In C. Fairburn (Ed.), *Cognitive behavior therapy and eating disorders* (pp. 265–308). New York: Guilford Press.
Fairburn, C. G., Cooper, Z., & Shafran, R. (2003). Cognitive behaviour therapy for eating disorders: a "transdiagnostic" theory and treatment. *Behaviour Research and Therapy, 41*, 509–528.
Fairburn, C. G., Jones, R., Peveler, R. C., Hope, R. A., & O'Connor, M. (1993). Psychotherapy and bulimia nervosa. Longer-term effects of interpersonal psychotherapy, behavior therapy, and cognitive behavior therapy. *Archives of General Psychiatry, 50*, 419–428.
Fairburn, C. G., Marcus, M. D., & Wilson, G. T. (1993). Cognitive-behavioral therapy for binge eating and bulimia nervosa: a comprehensive treatment manual. In C. Fairburn, & G. Wilson (Eds.), *Binge eating: Nature, assessment and treatment* (pp. 361–404). New York: Guilford Press.
First MB, RL, S., Gibbon, M., & Williams, J. (1997). User's guide for the Structured Clinical Interview for DSM-IV Axis I Disorders. Washington DC: American Psychiatric Press.
Insel, T. (2014). The NIMH research domain criteria (RDoC) project: precision medicine for psychiatry. *American Journal of Psychiatry, 171*, 395–397.
Insel, T., Cuthbert, B., Garvey, M., Heinssen, R., Pine, D. S., Quinn, K., et al. (2010). Research Domain Criteria (RDoC): toward a new classification framework for research on mental disorders. *American Journal of Psychiatry, 167*, 748–751.
Kazdin, A. (2003). Clinical significance: measuring whether interventions make a difference. In A. Kazdin (Ed.), *Methodological issues and strategies in clinical research* (3rd ed.). (pp. 691–710). Washington DC: American Psychological Association.
Kendall, P., Marrs-Garcia, A., Nath, S., & Sheldrick, R. (1999). Normative comparisons for the evaluation of clinical significance. *Journal of Consulting and Clinical Psychology, 67*, 285–299.
Klerman, G. L., Weissman, M. M., Rounsaville, B. J., & Chevron, E. S. (1984). *Interpersonal psychotherapy of depression*. New York: Basic Books.
Knott, S., Woodward, D., Hoefkens, A., & Limbert, C. (2015). Cognitive behaviour therapy for bulimia nervosa and eating disorders not otherwise specified: translation from randomized controlled trial to a clinical setting. *Behavioural and Cognitive Psychotherapy* (in press).
Loeb, K. L., Wilson, G. T., Labouvie, E., Pratt, E. M., Hayaki, J., Walsh, B. T., et al. (2005). Therapeutic alliance and treatment adherence in two interventions for bulimia nervosa: a study of process and outcome. *Journal of Consulting and Clinical Psychology, 73*, 1097–1107.
McIntosh, V. V. W., Jordan, J., Carter, F. A., Luty, S. E., McKenzie, J. M., Bulik, C. M., et al. (2005). Three psychotherapies for anorexia nervosa: a randomized, controlled trial. *American Journal of Psychiatry, 162*, 741–747.
Murphy, R., Cooper, Z., Hollon, S. D., & Fairburn, C. G. (2009). How do psychological treatments work? Investigating mediators of change. *Behaviour Research and Therapy, 47*, 1–5.
Murphy, R., Straebler, S., Basden, S., Cooper, Z., & Fairburn, C. G. (2012). Interpersonal psychotherapy for eating disorders. *Clinical Psychology & Psychotherapy, 19*, 150–158.
Poulsen, S., Lunn, S., Daniel, S., Folke, S., Mathiesen, B., Katznelson, H., et al. (2014). A randomized controlled trial of psychoanalytic psychotherapy or cognitive-behavioral therapy for bulimia nervosa. *American Journal of Psychiatry, 171*, 109–116.
Turner, H., Marshall, E., Stopa, L., & Waller, G. (2015). Cognitive-behavioural therapy for outpatients with eating disorders: effectiveness for a transdiagnostic group in a routine clinical setting. *Behaviour Research and Therapy, 68*, 70–75.
Wales, J., Palmer, R. L., & Fairburn, C. G. (2009). Can treatment trial samples be representative? *Behaviour Research and Therapy, 47*, 893–896.
Waller, G., Cordery, H., Corstorphine, E., Hinrichsen, H., Lawson, R., Mountford, V., et al. (2007). *Cognitive behavioral therapy for eating disorders*. Cambridge: Cambridge University Press.
Wilfley, D. E., Agras, W. S., Telch, C. F., EM, R., Schneider, J., Golomb, A., et al. (1993). Group cognitive-behavioral therapy and group interpersonal psychotherapy for the nonpurging bulimic individual: a controlled comparison. *Journal of Consulting and Clinical Psychology, 61*, 296–305.
Wilfley, D. E., Welch, R. R., Stein, R. I., Spurrell, E. B., Cohen, L. R., Saelens, B. E., et al. (2002). A randomized comparison of group cognitive-behavioral therapy and group interpersonal psychotherapy for the treatment of overweight individuals with binge-eating disorder. *Archives of General Psychiatry, 59*, 713–721.
Wilson, G. T., Wilfley, D. E., Agras, W. S., & Bryson, S. W. (2010). Psychological treatments of binge eating disorder. *Archives of General Psychiatry, 67*, 94–101.
Wonderlich, S., Peterson, C., Crosby, R., Smith, T., Klein, M., Mitchell, J., et al. (2014). A randomized controlled comparison of integrative cognitive-affective therapy (ICAT) and enhanced cognitive-behavioral therapy (CBT-E) for bulimia nervosa. *Psychological Medicine, 44*, 543–553.
Zipfel, S., Wild, B., Groß, G., Friederich, H., Teufel, M., Schellberg, D., et al. (2014). Focal psychodynamic therapy, cognitive behaviour therapy, and optimised treatment as usual in outpatients with anorexia nervosa (ANTOP study): randomised controlled trial. *The Lancet, 383*, 127–137.

Appendix 2: Mediation Trial Pre-Paper (Murphy et al., 2009)

Behaviour Research and Therapy 47 (2009) 1–5

Contents lists available at ScienceDirect

Behaviour Research and Therapy

journal homepage: www.elsevier.com/locate/brat

Invited Essay

How do psychological treatments work? Investigating mediators of change

Rebecca Murphy [a,*], Zafra Cooper [a], Steven D. Hollon [b], Christopher G. Fairburn [a]

[a] Department of Psychiatry, Oxford University, Warneford Hospital, Oxford OX3 7JX, UK
[b] Department of Psychology, Vanderbilt University, Nashville, TN 37240, USA

ARTICLE INFO

Article history:
Received 12 September 2008
Received in revised form 1 October 2008
Accepted 1 October 2008

Keywords:
Treatment
Mediation
Mechanisms
Interpersonal psychotherapy
Cognitive behaviour therapy
Eating disorders

ABSTRACT

Little is known about how psychological treatments work. Research on treatment-induced mediators of change may be of help in identifying potential causal mechanisms through which they operate. Outcome-focused randomised controlled trials provide an excellent opportunity for such work. However, certain conceptual and practical difficulties arise when studying psychological treatments, most especially deciding how best to conceptualise the treatment concerned and how to accommodate the fact that most psychological treatments are implemented flexibly. In this paper, these difficulties are discussed, and strategies and procedures for overcoming them are described.

© 2008 Elsevier Ltd. All rights reserved.

Introduction

While a large body of research has established the efficacy and effectiveness of a range of psychological treatments, little is known about how they work. This is an important shortcoming as most are limited in their efficacy. To help make them more potent, it would be of great value to understand how they work as research could then focus on enhancing the effective elements whilst discarding those elements found to be redundant (Kazdin & Nock, 2003). As matters stand markedly differing views are held regarding how psychological treatments work. For example, some claim that they work exclusively through common, or "non-specific", mechanisms (see Luborsky et al., 2002) a position that is hard to reconcile with the many studies that have identified treatment-specific effects (e.g., Clark et al., 2006; Fairburn, Jones, Peveler, Hope, & O'Connor, 1993). Opinions differ even when it comes to considering how a single psychological treatment works: for example, it has been argued that cognitive therapy for depression works by changing the content and structure of cognitive schema (Beck, Rush, Shaw, & Emery, 1979), through teaching compensatory skills (Barber & DeRubeis, 1989), or by establishing a metacognitive stance (Teasdale et al., 2002). To clarify such matters the processes responsible for treatment-induced change need to be identified and one step in this regard is the identification of mediators of change. What impedes such work, however, is the difficulty in designing and conducting mediational studies of psychological treatments. It is these difficulties, and the means of addressing them, that are the subject of this paper.

First, some concepts and terms need to be introduced. Mediators of treatment effects are variables which account for, in a statistical sense, at least some of the effects of treatment on the patient's outcome (Baron & Kenny, 1986). If the effects of treatment are mediated by a variable, this finding is consistent with the hypothesis that the treatment works by modifying this variable (the mediator). However, it must be stressed that mediation analysis involves merely identifying associations between putative variables: it does not establish that the identified relationship is causal in nature. The identification of mediators is, therefore, an initial step in establishing how treatments work, the next step being the testing of the causal status of any identified mediators by manipulating them. Thus, the value of identifying mediators lies in the narrowing down of the search for causal mechanisms.

Historically, there has been confusion and inconsistency over the use of the terms mediation and moderation. As defined by Kraemer, Wilson, Fairburn, and Agras (2002), moderators precede treatment, are uncorrelated with treatment and "explain", in a statistical sense, individual differences in the

* Corresponding author. Tel.: +44 1865 226479; fax: +44 1865 226244.
E-mail address: rebecca.murphy@psych.ox.ac.uk (R. Murphy).

0005-7967/$ – see front matter © 2008 Elsevier Ltd. All rights reserved.
doi:10.1016/j.brat.2008.10.001

effects of treatment. They indicate in whom and under what circumstances treatment has the most effect. In contrast, mediators are a consequence of treatment and "explain" in a statistical sense some of the effects of treatment on outcome. Thus, mediators correlate with treatment, are modified during treatment and this change precedes the effects of treatment on outcome.

Investigating mediators in the context of randomised controlled trials

Randomised controlled trials provide an often-missed opportunity to investigate the mediators of treatment effects and guidelines have been proposed for doing this (Kazdin, 2007; Kraemer et al., 2002). The key points are as follows. First, the decision to perform a mediator analysis needs to be taken in advance as it will influence the choice of measures used and when they are applied. Next, hypotheses need to be formulated concerning the likely mechanisms of action of the treatment under study. The hypotheses are likely to be derived from the theory underpinning the treatment and the findings of prior research. Then, the treatment, the putative mediators and the outcome need to be operationalised and a suitable assessment protocol devised. The ideal situation is when mediators are investigated in the context of trials that include a comparison treatment or control condition as this can help to rule out the possibility that what appears to mediate change is simply a general effect of receiving treatment (e.g., an expectancy effect) or a naturally occurring change (e.g., regression to the mean) rather than the specific effect of the treatment under consideration. The nature of the comparison treatment is also relevant. Minimal comparison treatments that are less effective than the index treatment can serve as controls for the processes just described, whereas treatments that are equally effective but operate via different mechanisms – as illustrated below – may be used to ensure that change in the putative mediator is not a consequence of change in the outcome variable ("reverse causality"). Note that if more than one treatment is being compared, it is possible to study the mediators of action of both treatments simultaneously, again as illustrated below.

The timing of the measurements is of critical importance as change in the mediator needs to be shown to have occurred prior to change in the outcome of interest as otherwise it could merely be a secondary effect. Establishing temporal precedence is particularly challenging when the effects of the mediator are likely to be rapid as is often the case. One solution is to measure both the putative mediator and the outcome variable at frequent intervals throughout the period over which change is likely to occur. Simultaneous measurement of both the mediator and the outcome variable also allows one to control for the level of the outcome when predicting its subsequent change. It is important to stress that it is not sufficient to simply measure the level of the outcome variable at some later point in time, for example, at the end of treatment. This is because treatment may produce change in the outcome prior to the point at which the mediator is assessed with the consequence that some portion of the change observed in the mediator could be the result of a causal path from early change in the outcome variable to subsequent change in the mediator (DeRubeis et al., 1990).

With regard to statistical analysis, a variety of data analytic strategies and procedures have been developed to assess whether a putative mediator meets statistical criteria for mediation. Readers are referred to the following sources (among others): Baron and Kenny (1986); MacKinnon, Lockwood, Hoffman, West, and Sheets (2002); Shrout and Bolger (2002) and MacKinnon, Fairchild, and Fritz (2007).

Studying the mediators of action of psychological treatments

Two particular challenges arise when studying the mediators of action of psychological treatments. First, psychological treatments differ in how they may be conceptualised. Some tend to be viewed as a complete unit rather than as a collection of procedures, whereas in others the opposite is the case. Interpersonal psychotherapy (IPT; Klerman, Weissman, Rounsaville, & Chevron, 1984; Weissman, Markowitz, & Klerman, 2000), for example, is usually conceptualised as a complete treatment or entity, as will be discussed below. In contrast, many cognitive behavioural treatments (CBT) comprise a variety of different procedural elements that may be viewed either in isolation or as an integrated group of techniques directed at a specific goal. For example, CBT for depression may be thought of either as a unit or as a treatment that comprises a number of distinct procedures including behavioural activation and cognitive restructuring (Jacobson et al., 1996). This distinction is important since it provides the possibility of either investigating the action of the treatment as a whole or that of its component parts.

The second matter is practical rather than conceptual. It is that all but the simplest psychological treatments are implemented in a flexible manner. In other words, their precise use is tailored to the individual patient's psychopathology, circumstances and progress at making change, and as a result the treatment is not identical in every case. For example, in most forms of CBT the precise time when particular procedures are implemented differs from patient to patient. Indeed, in some cases certain procedures may not be used at all. This flexibility, inherent to good clinical practice, substantially complicates research on mediation.

Investigating the mediators of action of CBT and IPT in the treatment of eating disorders

There follows a description of the strategies and procedures that we have developed to address certain of the difficulties specified above. The context is the treatment of patients with eating disorders but the principles apply across different disorders and different treatments. In this particular case, two treatments are being studied, interpersonal psychotherapy for eating disorders (Fairburn, 1997; Murphy, Straebler, Cooper, & Fairburn, in press) and "enhanced" transdiagnostic CBT for eating disorders (CBT-E; Fairburn, 2008a; Fairburn et al., in press). Both treatments are outpatient-based and involve 20 treatment sessions over 20 weeks. Initially the sessions are twice weekly, then after 4 weeks they are weekly for 10 weeks, and finally there are three sessions 2 weeks apart. Both treatments are the sole interventions that the patients receive other than continuation antidepressant medication in some cases.

IPT and CBT-E differ markedly in their rationale, strategies and procedures, and are very likely to have different modes of action. IPT is designed to help patients overcome their eating disorder indirectly by resolving current problems in their interpersonal life. The treatment is derived from IPT for depression (Klerman et al., 1984; Weissman et al., 2000) and closely resembles it. It has three phases. The first generally occupies 3–4 sessions. The aim is to describe the rationale and nature of the treatment and to identify jointly one or more current interpersonal problems that will thereafter become the focus of treatment. In the second phase, these problems are examined in detail with the therapist helping the patient first to characterise them and then identify possible means of addressing them. In the final phase, the focus shifts to the future, the goals being to ensure that any interpersonal changes made in treatment are maintained and to minimise the risk of relapse in the longer-term.

CBT-E is the latest version of the leading empirically supported treatment for eating disorders (Fairburn, 2008b). Originally, it was a treatment for bulimia nervosa but it is now transdiagnostic in nature; that is, it is designed to be suitable for all forms of eating disorder whatever the DSM-IV diagnosis. Unlike IPT, CBT-E comprises a collection of strategies and procedures focused directly on the characteristic psychopathology of eating disorders. While individual procedures target specific elements of this psychopathology (e.g., binge eating, body shape checking), the procedures are designed to be used in concert, the goal being to address all the main maintaining mechanisms operating in the individual patient. Thus, CBT-E can be conceptualised either as a collection of discrete procedures, each of which has its own mechanism of action, or as a complete unit. Complicating the investigation of CBT-E is the fact that it is designed to be used very flexibly with the choice of procedures and the timing of their implementation being tailored to the needs of the individual patient.

There follows a description of how we are attempting to identify the mediators of action of these two very different treatments starting with an outline of our hypotheses concerning how they might operate.

Hypothesised mediators of change in IPT

Little has been written about how IPT works. Drawing both on the original theory underpinning IPT (Klerman et al., 1984) and on our experience observing how patients change during treatment, we have formulated four hypotheses concerning the mechanisms through which IPT might achieve its effects. It is important to note that we do not view these as the sole mechanisms through which IPT operates nor are the hypotheses mutually exclusive. For example, in bulimia nervosa there is a rapid initial change in the frequency of binge eating and vomiting that is much the same in IPT and CBT (Fairburn, Agras, Walsh, Wilson, & Stice, 2004; Wilson, Fairburn, Agras, Walsh, & Kraemer, 2002). This is likely to be the result of common "non-specific" processes associated with starting treatment and, interestingly, its magnitude is a powerful predictor of outcome (Agras, Crow et al., 2000; Fairburn et al., 2004). Nevertheless, we hypothesize that in addition to these early non-specific processes (which operate in both IPT and CBT), IPT operates through four IPT-specific mechanisms.

1 *The reduction in eating disorder features is largely mediated by a decrease in the severity of the interpersonal problem(s) targeted in treatment.* In this case, the putative mediator is a reduction in the severity of the specific targeted interpersonal problem(s).
2 *The reduction in eating disorder features is largely mediated by an increase in interpersonal self-efficacy with regard to the specific problem(s) targeted in treatment.* Thus, in this case the putative mediator is an increase in the strength of the patient's belief that he or she is capable of overcoming the targeted interpersonal problem(s).
3 & 4 *The reduction in eating disorder features is largely mediated by an increase in general interpersonal self-efficacy (hypothesis 3) or self-esteem (hypothesis 4).* In the case of hypothesis 3, the proposed mediator is an increase in the strength of the patient's belief that he or she is capable of overcoming interpersonal difficulties in general rather than just those which have been the focus of treatment. In hypothesis 4, the putative mediator is an increase in self-esteem.

The nature of these four putative mediators is such that it is likely that they take considerable time to change. Therefore, in the case of all four hypotheses it is predicted that the critical change in the mediator takes place during the later stages of treatment and that the change in the eating disorder largely takes place subsequent to this. This prediction is consistent with the unexplained finding that patients with bulimia nervosa who receive IPT show significantly less change in their frequency of binge eating and vomiting than those who receive CBT, a difference that disappears over follow-up due to continuing improvement in the IPT patients and little or no change among those who received CBT (Agras, Walsh, Fairburn, Wilson, & Kraemer, 2000; Fairburn et al., 1993).

Hypothesised mediators of change in CBT-E

In the case of CBT-E, we have chosen to focus on the mediators of action of four specific treatment procedures rather than that of the treatment as a whole. Each procedure addresses a specific psychopathological process thought to be central to the maintenance of most eating disorders, and each is described in detail in the full treatment guide (Fairburn, 2008a).

1 The "weekly weighing" procedure – *It is predicted that the reduction in weight concern which occurs early on in CBT-E is largely mediated by a reduction in the frequency of weight checking (as a result of the "weekly weighing" procedure).* Patients with eating disorders are extremely concerned about their weight (i.e., the number on the scale). They often check their weight very frequently and as a result become focused on inconsequential weight changes (Fairburn, 2008b). This over-concern with weight, and the associated fear of weight gain, maintains strict dietary restraint and is a barrier to patients changing their way of eating. It is addressed early on in CBT-E by the weekly weighing procedure in which the therapist and patient jointly check the patient's weight once a week and then together interpret the resulting finding in the light of education and prior readings. Patients are helped not to weigh themselves between these times.
2 The "regular eating" procedure – *It is predicted that the reduction in the frequency of binge eating which occurs early on in CBT-E is largely mediated by the adoption of a pattern of regular eating (as a result of the "regular eating" procedure).* Patients with eating disorders have a distinctive way of eating. Their temporal pattern of eating tends to be characterised by delayed eating and the avoidance of meals or snacks, or by a highly unstructured way of eating (Fairburn, 2008b). Regular eating, like weekly weighing, is introduced early on in treatment. It involves helping patients establish a daily pattern of eating characterised by three planned meals and two snacks with no eating in between. Regular eating of this type has a variety of effects one of which is to displace binge eating.
3 The "dietary rules" procedure – *The reduction in dietary restraint seen in the mid-to-late stages of CBT-E is largely mediated by erosion of the belief that adherence to strict dietary rules is necessary to prevent binge eating, weight gain or fatness (as a result of the "dietary rules" procedure).* The great majority of patients with eating disorders engage in a highly distinctive form of dieting characterised by attempts to adhere to multiple strict and demanding dietary rules (Fairburn, 2008b). Doing so is valued by the patients. The "dietary rules" procedure is designed to address this form of dieting. It involves identifying patients' dietary rules and highlighting the fact that they substantially impair their day-to-day life and maintain their eating problem, and so need to be targeted in treatment. Patients' fears about breaking these rules are identified and challenged using behavioural experiments in combination with relevant education about the triggers of binge eating and the processes underlying weight gain and shape change. As a result, patients learn that the feared consequences of breaking these rules do not occur and consequently the rules become viewed as a problem rather than as a strength. This erosion of the belief

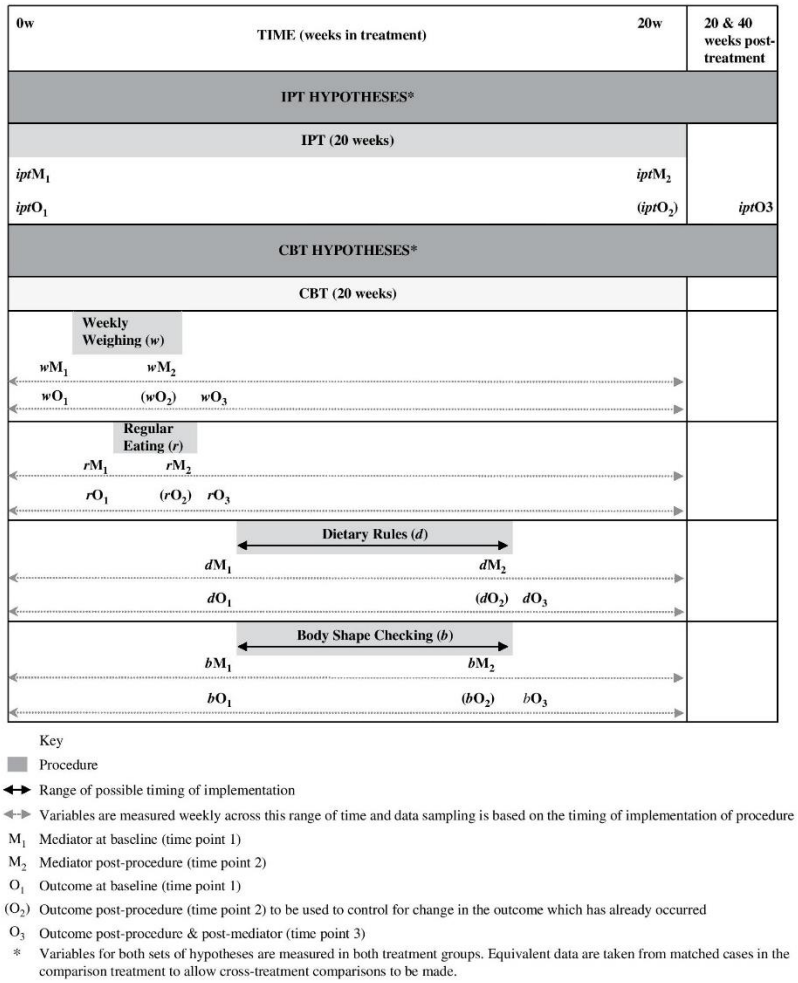


Fig. 1. Schematic diagram illustrating the research design used to identify mediators of action of IPT and CBT-E in the treatment of patients with eating disorders.

that adherence to strict dietary rules is necessary to prevent binge eating, weight gain or fatness leads patients to moderate and eventually abandon their dietary restraint. The use and timing of this intervention is tailored to the needs of the individual patient based on an evaluation of the importance of "dietary rules" in maintaining whatever eating disorder psychopathology remains after the first stage of treatment.

4 The "body shape checking" procedure – *The reduction in concern about body shape which occurs in the mid-to-late stages of CBT-E is largely mediated by a decrease in body shape checking (as a result of the "body shape checking" procedure).* Patients with eating disorders typically engage in frequent shape checking (using multiple methods including body pinching, body measuring, and scrutinising themselves in mirrors) (Fairburn, 2008b). In the "body shape checking" procedure, patients are first made aware of their body checking (much of which may occur outside their awareness) and its adverse effects. Patients are then helped to make strategic changes to their shape checking and to evaluate the significance of the resulting effects. This results in a decrease in their concerns about body shape. In common with the dietary rules procedure, the use and timing of this procedure is personalised to the individual patient based on the importance of shape concern in maintaining their remaining eating disorder psychopathology.

In contrast with IPT's hypothesised mechanisms of action, the four CBT procedures are thought to operate rapidly with there being a short time lag between change in the putative mediator and change in the relevant outcome variable.

Research design

In terms of research design, the investigation of the four putative mediators of action of IPT is relatively straightforward (and is illustrated in Fig. 1). The main points are as follows:

1 IPT is conceptualised as a unit;
2 the mediator is measured at baseline and at the end of IPT (i.e., 20 weeks later);
3 the outcome of interest (level of eating disorder psychopathology) is measured at the outset of IPT and 20 and 40 weeks after the end of treatment (i.e., well after the end-of-IPT measurement of change in the putative mediator, thereby addressing temporal precedence). The level of eating disorder psychopathology is reassessed at the end of treatment so that any incidental in-treatment change may be controlled for in the analysis;
4 equivalent data are taken from individually matched CBT-E cases to allow cross-treatment comparisons to be made.

The same principles are used to address the four CBT-E hypotheses. However, in this case there are two additional complexities. The first is that change in the mediator is thought to occur rapidly and as a result there may be a relatively short interval before change in the outcome occurs. The second is that the third and fourth of the CBT-E procedures under study are not introduced at set points in treatment. As a result, the design has to be modified in four ways (Fig. 1):

1 the timing of implementation of each of the four CBT-E procedures is recorded on a case-by-case basis so that it is known exactly when they took place;
2 each mediator and outcome variable is measured frequently and simultaneously throughout treatment;
3 the data needed for the mediational analyses are selected on a patient-by-patient basis depending on when each of the four procedures was implemented. In each case, the data on the level of the mediator is extracted just before the implementation of the procedure concerned, and also at the end of its implementation. Similarly, data on the level of the outcome variable is extracted just prior to the onset of the implementation of the procedure and shortly after the reassessment of the level of the mediator. Thus, in this way a "floating", but predetermined, individualised dataset is extracted for analysis;
4 equivalent data are collected from an individually matched IPT case so that cross-treatment, temporally matched, comparisons may be made.

Conclusions

In this paper, we have discussed in principle how research on mediators of change may be of help in identifying potential causal mechanisms through which psychological treatments operate. Following Kraemer et al. (2002), we have argued that such work can fruitfully be combined with outcome-focused randomised controlled trials. We have noted that certain conceptual and practical challenges arise when studying psychological treatments, two being particularly important. The first is how best to conceptualise the treatment under investigation, the core decision being whether to view it as a unit or as a collection of procedures each with its own mode of action. The second stems from the fact that most psychological treatments are implemented flexibly. We have discussed potential ways of addressing these problems based on our work evaluating the mediators of action of IPT and CBT-E in the treatment of patients with eating disorders. The same principles should apply to the investigation of other psychological treatments and other psychiatric disorders.

Acknowledgements

We are grateful to the Wellcome Trust for its support. RM and ZC are supported by a programme grant from the Trust (046386), and CGF is supported by a Principal Research Fellowship (046386). We are grateful to Eleanor Frampton-Fell for her help preparing the manuscript.

References

Agras, W. S., Crow, S. J., Halmi, K. A., Mitchell, J. E., Wilson, G. T., & Kraemer, H. C. (2000). Outcome predictors for the cognitive behavior treatment of bulimia nervosa: data from a multisite study. *American Journal of Psychiatry, 157*, 1302–1308.

Agras, W. S., Walsh, T., Fairburn, C. G., Wilson, G. T., & Kraemer, H. C. (2000). A multicenter comparison of cognitive-behavioral therapy and interpersonal psychotherapy for bulimia nervosa. *Archives of General Psychiatry, 57*, 459–466.

Barber, J. P., & DeRubeis, R. J. (1989). On second thought: where the action is in cognitive therapy for depression. *Cognitive Therapy and Research, 13*, 441–457.

Baron, R. M., & Kenny, D. A. (1986). The moderator-mediator variable distinction in social psychological research: conceptual, strategic and statistical considerations. *Journal of Personality and Social Psychology, 51*, 1173–1182.

Beck, A. T., Rush, A. J., Shaw, B. F., & Emery, G. (1979). *Cognitive therapy of depression*. New York: Guilford.

Clark, D. M., Ehlers, A., Hackmann, A., McManus, F., Fennell, M., Grey, N., et al. (2006). Cognitive therapy versus exposure and applied relaxation in social phobia: a randomized controlled trial. *Journal of Consulting and Clinical Psychology, 74*, 568–578.

DeRubeis, R. J., Evans, M. D., Hollon, S. D., Garvey, M. J., Grove, W. M., & Tuason, V. B. (1990). How does cognitive therapy work? Cognitive change and symptom change in cognitive therapy and pharmacotherapy for depression. *Journal of Consulting and Clinical Psychology, 58*, 862–869.

Fairburn, C. G. (1997). Interpersonal psychotherapy for bulimia nervosa. In D. M. Garner, & P. E. Garfinkel (Eds.), *Handbook of treatment for eating disorders* (2nd ed.). (pp. 278–294) New York: Guilford.

Fairburn, C. G. (2008a). *Cognitive behavior therapy and eating disorders*. New York: Guilford Press.

Fairburn, C. G. (2008b). Eating disorders: the transdiagnostic view and the cognitive behavioral theory. In C. G. Fairburn (Ed.), *Cognitive behavior therapy and eating disorders* (pp. 7–22). New York: Guilford Press.

Fairburn, C. G., Agras, W. S., Walsh, B. T., Wilson, G. T., & Stice, E. (2004). Prediction of outcome in bulimia nervosa by early change in treatment. *American Journal of Psychiatry, 161*, 2322–2324.

Fairburn, C. G., Cooper, Z., Doll, H. A., O'Connor, M. E., Bohn, K., Hawker, D. M., et al. Transdiagnostic cognitive behavior therapy for patients with eating disorders: a two-site trial with 60-week follow-up. *American Journal of Psychiatry*, in press.

Fairburn, C. G., Jones, R., Peveler, R., Hope, R., & O'Connor, M. (1993). Psychotherapy and bulimia nervosa: longer-term effects of interpersonal psychotherapy, behavior therapy, and cognitive behavior therapy. *Archives of General Psychiatry, 50*, 419–428.

Jacobson, N. S., Dobson, K. S., Truax, P. A., Addis, M. E., Koerner, K., Gollan, J. K., et al. (1996). A component analysis of cognitive-behavioral treatment for depression. *Journal of Consulting and Clinical Psychology, 64*, 295–304.

Kazdin, A. E. (2007). Mediators and mechanisms of change in psychotherapy research. *Annual Review of Clinical Psychology, 3*, 1–27.

Kazdin, A. E., & Nock, M. K. (2003). Delineating mechanisms of change in child and adolescent therapy: methodological issues and research recommendations. *Journal of Child Psychology and Psychiatry, 44*, 1116–1129.

Klerman, G. L., Weissman, M. M., Rounsaville, B. J., & Chevron, E. S. (1984). *Interpersonal psychotherapy for depression*. New York: Basic Books.

Kraemer, H. C., Wilson, G. T., Fairburn, C. G., & Agras, W. S. (2002). Mediators and moderators of treatment effects in randomized clinical trials. *Archive of General Psychiatry, 59*, 877–883.

Luborsky, L., Rosenthal, R., Diguer, L., Andrusyna, T. P., Berman, J. S., Levitt, J. T., et al. (2002). The dodo bird verdict is alive and well—mostly. *Clinical Psychology, 9*, 2–12.

MacKinnon, D. P., Fairchild, A. J., & Fritz, M. S. (2007). Mediation analysis. *Annual Review of Psychology, 58*, 593–614.

MacKinnon, D. P., Lockwood, C. M., Hoffman, J. M., West, S. G., & Sheets, V. (2002). A comparison of methods to test mediation and other intervening variable effects. *Psychological Methods, 7*, 83–104.

Murphy, R., Straebler, S., Cooper, Z., & Fairburn, C. G. (in press). Interpersonal psychotherapy (IPT) for eating disorders. In Dancyger, I. & Fornari, V. (Eds.), *Evidence based treatments for eating disorders: Children, adolescents, and adults*. New York: Nova Science Publishers.

Shrout, P. E., & Bolger, N. (2002). Mediation in experimental and nonexperimental studies: new procedures and recommendations. *Psychological Methods, 7*, 422–445.

Teasdale, J. D., Moore, R. G., Hayhurst, H., Pope, M., Williams, S., & Segal, Z. V. (2002). Metacognitive awareness and prevention of relapse in depression: empirical evidence. *Journal of Consulting and Clinical Psychology, 70*, 275–287.

Weissman, M. M., Markowitz, J. C., & Klerman, G. L. (2000). *Comprehensive guide to interpersonal psychotherapy*. New York: Basic Books.

Wilson, G. T., Fairburn, C. G., Agras, W. S., Walsh, B. T., & Kraemer, H. C. (2002). Cognitive-behavioral therapy for bulimia nervosa: time course and mechanisms of change. *Journal of Consulting and Clinical Psychology, 70*, 267–274.

Study Measures

Appendix 3: Regular Eating Rating Scale (RERS)

REGULAR EATING RATING SCALE

First, check that a sufficient proportion of the available records are present in order to make a valid rating.
This rating is based on adherence to the CBT-E "regular eating" procedure, as described in the manual, over the duration of one week of treatment.
Successful adherence involves two components:

1. Positive features of occasions of eating at least every four hours (between the meals and snacks or after waking/before sleeping using common sense approach as these times are unlikely to be unspecified)
    - Meals and snacks should not be followed by vomiting or laxative/diuretic misuse.
    - Note that binge eating episodes do not count as meals or snacks.
2. If positive features present also need check for gaps between eating occasions
    - Note that binge eating episodes do not count here as meals or snacks – someone might have marked adherence to regular eating whilst having episodes of binge eating.
    - For most people this will mean eating is largely confined to three planned meals each day, plus two or three planned snacks.

If in doubt, rate up i.e. cautiously on the side of less pathology

Global Rating Scheme

In general rate the most common pattern of eating throughout the week using the rating scheme below. If it is difficult to choose between two ratings, the more conservative rating (i.e., higher) should be made.

- 6 – Marked adherence to regular eating every day (but with minor slips allowed). For example, eating on 5 or 6 occasions each day with roughly evenly spaced gaps in between.
- 5 – Pretty good adherence but not full adherence. For example, when full adherence is present on the majority of the days with the occasional too-long or too-short gaps on the other days.
- 4 – Moderate adherence. For example, when full adherence occurs most days (4 or 5 days) with some too-long or too-short gaps on the other days. Or perhaps consistent too-long or too-short gaps.
- 3 – Some adherence. When the patient eats regularly at times but at other times has too long / too short gaps between eating episodes. For example, when regular eating happens just less than half of the days or only half of the time.
- 2 – Slight adherence. Occasionally eating regularly.
- 1– Rarely eating regularly (i.e., marginal adherence). For example for only one day a week (where seven days are rated) or the occasional part of the day.
- 0 – No regular eating. For example, either long periods of time without eating, or only binge eating or grazing throughout day. No days or parts of days when regular eating occurs.

Appendix 4: Regular Eating Rating Scale (RERS) and Binge Eating Rating Scale (BERS) Rating Form

Rating Form for the 'Measure of Regular Eating' and 'Measure of Objective Binge Eating'

1. Date: …….
2. Trial Number: ……
3. Treatment Week: ……
4. Gender: Male/Female (please circle)
5. Number of records present: ……..
6. Number of records that should have been completed: ……..
7. Missing records: ……..
8. Sufficient number of records to make a rating: Y / N (please circle)

Instructions

The rater should:

A. Fill in the 'date'
B. Note the 'patient trial number' will already have been entered.
C. Note the 'treatment week' will already have been blindly coded and entered.
D. Note the 'gender' of the patient will have already have been circled.
E. Fill in the total 'number of records present'.
F. Note the 'number of records that should have been completed' will already have been entered.
G. Note the records will have been numbered per day and any missing records entered above.
H. To obtain a rating there needs to be a good proportion of records available, from those that should have been completed (see the 'Measure of Regular Eating' within the research protocol.
I. To rate regular eating, you will need to look at the 'Measure of Regular Eating' within the research protocol (pg. 3). Using the global rating scheme for regular eating, enter the rating into the box below. You can also write any 'notes' that you require.
J. To rate objective binge eating episodes, you will need to look at the 'Measure of Objective Binge Eating' within the research protocol (pg. 4). To the right of the rating form there is space to enter the amount of OBEs for each monitoring record per day. Calculate and enter the mean number of binges per day into the box at the bottom of the form. This need not be an integer and should be to one decimal place.
    - When looking at the monitoring records, note that some of the records are across 2-3 pages. Look carefully at each monitoring record to ensure that no information is missed.

"Measure of Regular Eating"

Notes:

9. Rating: Most common pattern of eating:

"Measure of Objective Binge Eating"

Number of binges – record 1 ……..
Number of binges – record 2 ……..
Number of binges – record 3 ……..
Number of binges – record 4 ……..
Number of binges – record 5 ……..
Number of binges – record 6 ……..
Number of binges – record 7 ……..
Number of binges – record 8 ……..
Number of binges – record 9 ……..
Number of binges – record 10 ……..
Number of binges – record 11 ……..
Number of binges – record 12 ……..
Number of binges – record 13 ……..
Number of binges – record 14 ……..

10. Rating:
Mean number of binges per day:

Rating Form Discussed: Y/N | Final Regular Eating Rating: | Final Objective Binge Eating Rating:

Appendix 5: Binge Eating Rating Scale

OBJECTIVE BINGE EATING RATING SCALE

This rating is based on the frequency of objective binges (OBEs) on average for each day over a week of treatment.

An objective binge comprises two criteria:

1. Eating, in a discrete period of time (e.g., within any 2-hour period), an amount of food that is definitely larger than most people would eat in a similar period of time under similar circumstances
2. Accompanied by a sense of loss of control.

However, the rater does not have information on loss of control, so one is largely rating "objective overeating" as a proxy for objective binge eating. Note that it is not necessary for episodes of overeating to be followed by asterisks or purging in order to make a rating of an OBE.

Rating Scheme

The rater will total the number of OBEs for each day and calculate a mean per day. If it is unclear whether or not it is a binge, then if in doubt it should not count as one and a conservative rating (i.e., lower) should be made.

Rate each episode of OBE according to each occasion of objective overeating:

- This decision should be made by the rater. To rate OBEs, an unusually large amount of food needs to have been consumed. A large amount of food can refer to either; the amount of any particular type of food, or the overall quantity of food consumed. Either way, the amount should be clearly large, although it does not need to be enormous for this to be rated. Therefore the rater will need to use their judgement, alongside any other information present to make the decision about whether or not objective overeating occurred. In deciding whether the amount was "large", the rater can consider the other circumstantial information available i.e. time of day or whether it has been identified as a 'meal'. It is also important that the rater considers the patient's sex as males in general would be expected to consume larger quantities of food then females.
- When calculating the number of episodes of overeating, one should use a significant time gap when the participant was not eating to signify the completion of an episode of overeating. Note that "purging" (self-induced vomiting or laxative misuse) is not used to define the end of individual episodes of overeating.

Appendix 6: Eating Questionnaire (EQ)

EATING QUESTIONNAIRE

Instructions: The following questions are concerned with the past week (7 days). Please read each question carefully. Please answer all of the questions. Thank you.

Questions 1 to 12: Please circle the appropriate NUMBER OF DAYS on the right. Remember that the questions only refer to the past week (7 days).

	On how many of the past 7 days …..	0 days	1 day	2 days	3 days	4 days	5 days	6 days	7 days
1.	Have you been deliberately trying to limit the amount of food you eat to influence your shape or weight (whether or not you have succeeded)?	0	1	2	3	4	5	6	7
2.	Have you gone for long periods of time (8 waking hours or more) without eating anything at all in order to influence your shape or weight?	0	1	2	3	4	5	6	7
3.	Have you tried to exclude from your diet any foods that you like in order to influence your shape or weight (whether or not you have succeeded)?	0	1	2	3	4	5	6	7
4.	Have you tried to follow definite rules regarding your eating (for example, a calorie limit) whether or not you have succeeded?	0	1	2	3	4	5	6	7
5.	Have you had a definite desire to have an empty stomach with the aim of influencing your shape or weight?	0	1	2	3	4	5	6	7
6.	Have you had a definite desire to have a totally flat stomach?	0	1	2	3	4	5	6	7
7.	Has thinking about food, eating or calories made it very difficult to concentrate on things you are interested in (for example, working, following a conversation, or reading)?	0	1	2	3	4	5	6	7
8.	Has thinking about shape or weight made it very difficult to concentrate on things you are interested in (for example, working, following a conversation, or reading)?	0	1	2	3	4	5	6	7
9.	Have you had a definite fear of losing control over eating?	0	1	2	3	4	5	6	7
10.	Have you had a definite fear that you might gain weight?	0	1	2	3	4	5	6	7
11.	Have you felt fat?	0	1	2	3	4	5	6	7
12.	Have you had a strong desire to lose weight?	0	1	2	3	4	5	6	7

Questions 13-16: Please fill in the appropriate NUMBER OF TIMES in the spaces on the right. Remember that the questions only refer to the past week (7 days).

Over the past week (7 days) ……

Over the past 7 days, how many times have you eaten an unusually large amount of food (given the circumstances) and had a sense of having lost control over eating? ……..

Over the past 7 days, how many times have you made yourself sick (vomited) as a means of controlling your shape or weight? ……..

15a. Over the past 7 days, how many times have you taken laxatives as a means of controlling your shape or weight? ……..

15b. Over the past 7 days, have you engaged in any other means of controlling shape or weight (e.g. taking any other tablets)? If yes, please give details of this below, including the number of times you have engaged in this behavior:

……

16. Over the past 7 days, how many times have you weighed yourself not including in-session weighing? ……..

Questions 17 to 21: Please circle the appropriate response. Please note that for the remaining questions the term "binge eating" means eating what others would regard as an unusually large amount of food for the social circumstances, accompanied by a sense of having lost control over eating.

Over the past week (7 days) ……

17. Over the past 7 days, on how many days have you eaten in secret (i.e., furtively)? ….. Do not count episodes of binge eating	0 days	1 day	2 days	3 days	4 days	5 days	6 or 7 days

18. Over the past 7 days, how often have you *actively* checked your body shape, size, or parts of your body (e.g. stomach or thighs) in a mirror?	Not at all	A few times	Once a day	A few times a day	Many times a day
19. Over the past 7 days, how often have you pinched or measured your body (or parts of your body)?	Not at all	A few times	Once a day	A few times a day	Many times a day
20. Over the past 7 days how often have you compared your body or parts of your body to others?	Not at all	A few times	Once a day	A few times a day	Many times a day
21. Over the past 7 days, on what proportion of the times that you have eaten have you felt guilty (felt that you've done wrong) because of its effect on your shape or weight?	None of the times	Some of the time	Half of the times	Most of the time	Every time

Questions 22 to 30: Please circle the appropriate number on the right. Remember that the questions only refer to the past week (7 days).

Over the past 7 days …..	Not at all		Slightly		Moderately		Markedly
22. How concerned have you been about other people seeing you eat? ….. Do not count episodes of binge eating	0	1	2	3	4	5	6
23. Has your weight influenced how you think about (judge) yourself as a person?	0	1	2	3	4	5	6
24. Has your shape influenced how you think about (judge) yourself as a person?	0	1	2	3	4	5	6
25. How much would it have upset you if you had been asked to weigh yourself once a week (no more, or less, often) for the next four weeks?	0	1	2	3	4	5	6
26. How worried have you been about your weight (i.e. the number on the scale)?	0	1	2	3	4	5	6
27. How worried have you been about your body shape (e.g. stomach, thighs)?	0	1	2	3	4	5	6
28. How uncomfortable have you felt seeing your body (for example, seeing your shape in the mirror, in a shop window reflection, while undressing or taking a bath or shower)?	0	1	2	3	4	5	6
29. How uncomfortable have you felt about others seeing your shape or figure (for example, in communal changing rooms, when swimming, or wearing tight clothes)?	0	1	2	3	4	5	6
30. To what extent have you believed that it is necessary to follow "rules" about eating in order to prevent weight gain, fatness or binge eating?	0	1	2	3	4	5	6
THANK YOU							

Appendix 7: Interpersonal Questionnaire (IQ)

INTERPERSONAL QUESTIONNAIRE

This questionnaire is concerned with your relationships, life and how you have been feeling. Please circle the number on the right which best describes how you have been OVER THE PAST WEEK (7 days). Thank you.

Over the past WEEK on AVERAGE	Not at all		A little		Quite a bit		Very much so
Have you been happy with your relationships?	0	1	2	3	4	5	6
Have you had arguments or disagreements with important people in your life?	0	1	2	3	4	5	6
Have you wanted more friends?	0	1	2	3	4	5	6
Have you felt close to anyone?	0	1	2	3	4	5	6
Have you felt that you have had someone you could share your feelings with (if needed)?	0	1	2	3	4	5	6
Have you felt you couldn't get on with people?	0	1	2	3	4	5	6
Have you felt lonely?	0	1	2	3	4	5	6
Have you had a good relationship with your partner? (leave blank if n/a)	0	1	2	3	4	5	6
Have you been thinking about someone close to you who has died?	0	1	2	3	4	5	6
Have you got on well with your family?	0	1	2	3	4	5	6
Have you been having difficulties adjusting to a change in your life? (e.g. marriage/divorce, job change, moving) (leave blank if no changes)	0	1	2	3	4	5	6
Have you been able to be assertive in your relationships?	0	1	2	3	4	5	6
Have you felt capable of overcoming or dealing with difficulties in your personal life?	0	1	2	3	4	5	6
Have you felt lacking in self-confidence?	0	1	2	3	4	5	6
Have you been feeling sad?	0	1	2	3	4	5	6
Have you lost interest in things?	0	1	2	3	4	5	6
Have you felt like a worthwhile person?	0	1	2	3	4	5	6
Have you been feeling guilty?	0	1	2	3	4	5	6
Have you been able to enjoy things as much as usual?	0	1	2	3	4	5	6
Have you disliked yourself or been self-critical?	0	1	2	3	4	5	6
Have you felt hopeful about the future?	0	1	2	3	4	5	6

Appendix 8: Interpersonal Interview (IPI)

Interpersonal Interview Schedule

INTERVIEW DETAILS
[These are to be completed before the interview using the Interpersonal Problem Form for details of the interpersonal problem(s).]

Date: [ ][ ]-[ ][ ]-[ ][ ]
Interpersonal Problem(s):

Complete using the participant's own words, including (if applicable) the name or role of any other parties and tick (check) the interpersonal problem area(s) to which the problem belongs. Note that it is possible for an interpersonal problem to fulfil the criteria for more than one IPT problem area.

(i) Primary Problem ……
(ii) Secondary Problem ……..
(iii)Tertiary Problem ……..

Problem	Interpersonal Problem Area(s)				
	Differing Expectations (Interpersonal Conflict)	Interpersonal Deficit: Narrow or Broad *	Role Transition	Grief	Life Goals
(i)Primary					
(ii)Secondary					
(iii)Tertiary					

* This interpersonal problem area includes anyone who appears to have a general lack of relationships with other people. It also includes anyone who appears to have a difficulty with being able to get close to people with whom relationships already exist or having concerns with regards to an intimate relationship e.g. wanting an intimate relationship or being able to keep it going.

INTRODUCTION
Firstly, I would like to remind you to please not tell me anything about your treatment.
You and your therapist have discussed difficulties in your life and relationships and how they affect you. (At the start of treatment) You arrived at the decision that [insert problem/s] is/are something which affects you and is/are a current difficulty in your life. I am going to ask you some questions about this. This is part of our research. So for this meeting we are only going to focus on [insert problem/s], (as it was defined then) even though you may discuss lots of other important things with your therapist.
As this is a partially structured interview whereby a standard set of questions is going to be asked, some might not apply to you.
I would like to ask you some questions about [insert problem/s] and I would like you to answer them….

[If at the start of treatment]… focusing on the 4 weeks before treatment begun.

[If at end of treatment or follow-up] … focusing on the past 4 weeks.

To help you to recall this time period I have made a calendar that specifies it.

[Show blank calendar. However, if there is already a calendar for this time period from another research interview this may be used instead]

So it goes from…
[If at the start of treatment]…the day before treatment started [day and date] back to [day and date]
OR
[If at the end of treatment or follow-up] …yesterday [day and date] back to [day and date]

What I would like you to do now is tell me about any events that have happened in these four weeks since this will help us discuss them. Have there been any events out of the ordinary ..... such as celebrations of any type, trips away or days off work? Then we can note these down on the calendar.

[These should be noted on the calendar thereby allowing the interviewer and participant to use it as an aide memoire.]

[Now read the questions which follow aloud to the patient. Rate the patient's responses to the following questions using the rating scheme below. Remind the patient that all questions ask about how things have been on average over the last four weeks / four weeks before treatment.]

1. SEVERITY OF INTEPERSONAL PROBLEM

[All questions refer to the problem specified on the 'Therapist Problem Form' (see Appendix A). For interpersonal interviews conducted after the initial interview the interviewer should rate the same problem as that identified at baseline. Question 1 should be asked for each problem and they should be rated separately. If the individual responds with a closed answer then it may be appropriate to ask 'Could you tell me more about that?']

General description of the interpersonal problem
Could you tell me about [insert problem/s]?
Has this been a difficulty for you over the last four weeks/four weeks before treatment?
Has [insert problem/s] been something you would like to change?
Have there been any difficulties associated with [insert problem/s]?
How much of a problem/How bad/How difficult do you think [insert problem/s] is, based on how it has been over the last four weeks/four weeks before treatment?
Does it affect you only at the time [e.g. that conflict is happening] or at other times too?

[Note that interviewers do not need *precise* details of the problem and how it affects the participant's life. However, interviewers require enough information to be able to make a rating.
The interviewer should rate the state of the interpersonal problem independently of the participant's beliefs about its cause (e.g. the eating disorder) in order to rate how present the problem is. However, the interviewer should find out about difficulties that the participant believes are caused by, or associated with, the interpersonal problem in order to assess impact.
If there are two problems then it is best to ask questions related to each in turn (i.e. question 1 for the first interpersonal problem followed by question 2 as regards this same interpersonal problem, followed by questions 1 and 2 for the second problem, ending with question 3.)
The 'problem' as defined at baseline remains the topic of interest. The patient may no longer agree with the previous conceptualisation of the problem ("actually, looking back, I don't think that was a problem") but the rating is of the current severity of the previously defined problem.]

Specific aspects of the interpersonal problem
Now ask questions specific to the problem area. The therapist should point you towards any section that they feel may be helpful. However, you should also use your judgement to use questions from any 'problem area' that seems relevant. Questions not included on the schedule can also be asked (any questions you wish to devise which seem relevant) which would help you to further understand

the nature of the difficulty. Of note, if the problem is initially 'differing expectations' and the relationship ends then selected 'role transition' questions may be relevant. – See Checklist

Impairment across a range of domains due to the interpersonal problem
[Ask about this over the four weeks]
I'm now going to ask you about how [insert problem/s] affects you and your life.
Does [insert problem/s] affect you're eating? [If negative effect]: In what way? How much?
Does [insert problem/s] affect how you feel / your mood? [If negative effect]: In what way? How much?
Does [insert problem/s] affect your self-confidence or how you feel about yourself as a person? [If negative effect]: In what way? How much?
Does [insert problem/s] affect how your work? [If negative effect]: In what way? How much?
Does [insert problem/s] affect other (aspects of your) relationships (or socializing)? [If negative effect]: In what way? How much?
Does [insert problem/s] affect your physical health (e.g. feeling tired)? [If negative effect]: In what way? How much?
Does [insert problem/s] affect anything else? [If negative effect]: In what way? How much?
Is there anything [insert problem/s] stops you from doing or gets in the way of? [If negative effect]: In what way? How much?
[If it seems as if there is no impairment this could be checked by asking this question in a summarised form: I'm now going to ask you about how [insert problem/s] affects you and your life. Does [insert problem/s] cause you any problems in terms of having a negative impact on your eating, mood, self-confidence, work, relationships, physical health and anything else?]

Rating Scale for the Severity of the Interpersonal Problem

Severity ratings should reflect a combination of two inter-related factors:
(I) How *present* the problem is (the frequency and duration of the difficulties) and
(ii) How much *impact* the problem has on the participant (impairment associated with the Problem).

The rating should reflect the interviewer's judgement of the severity of the problem based on the participant's responses. This is distinct from the participant's view of how severe the problem is; although in many cases the two will be similar.

The rating of the severity of the problem is based on its status over the past one month. Relative differences between the problem now and at another time-point are not of relevance to the rating. Even if the relationship in question has ended, the same problem is rated.

0 – No problem / problem resolved - Either no difficulties throughout the four weeks or a rare or minor difficulty. No impairment.
1 – Very infrequent difficulties and only of short duration. There might be very slight impairment.
2 – Slight problem, not significantly interfering with life
Difficulties occurring occasionally during the four weeks but they are not enduring. There may be slight impairment.
3 – Difficulties occurring on quite a few occasions over the four weeks. Difficulties somewhat more prolonged. Some degree of limited impairment.
4 – Moderate problem, somewhat interfering with life
Longer-lasting difficulties occurring quite often over the four weeks. Some degree of impairment across a range of domains.
5 – Difficulties occurring several times a week over the four weeks and these are long-lasting. Significant impairment.
6 – Marked problem, significantly interfering with life. Difficulties occurring most of the time and they have resulted in substantial impairment.
1(i) [ ]

1(ii) [ ]
1(iii) [ ]

Checklist for Assessing Severity of Interpersonal Problems
[Below are some questions that might assist in assessing interpersonal problems and their severity. They should be chosen and adapted as appropriate. It is preferable to use open rather than closed questions. Note that it is worth continuously checking that this applies to the period of time of interest (four weeks prior to treatment or previous four weeks), perhaps by using the calendar or asking e.g. 'was that also true for the four weeks...'. Even if the question seems phrased in a more general way (e.g. is X a problem?) it is written in this way to avoid repetition and it is referring to the four weeks specifically (e.g. was X a problem in the four weeks). If the individual responds with a closed answer or doesn't give enough information then it may be appropriate to ask 'Could you tell me more about that?'].

Differing role expectations (four weeks prior to treatment or previous four weeks).
Have you been having any difficulties within the relationship with [insert identified person]? Can you tell me about them?
Are you having disagreements with [insert identified person]?[If yes]
How often are you having disagreements?
How bad does it get? What happens when you disagree? Do you ever end up in physical conflict?
Are there things that are difficult to resolve in your relationship? Things that you do not agree on?
Do you each have different expectations of each other, or different things that you want, or different ideas about how you would like things to be?
Are you able to tell ……. how you feel about [the area of disagreement / area of differing expectations]?
Are you able to help [the other person] understand your point of view? Are you able to see their point of view?
Are you able to resolve your disagreements / differing expectations?
[If yes]
How does this happen?
How long does it take?
Is there a difficult atmosphere between the two of you? How long does this last?
Have you stopped talking? Are you silent with each other?
Have you avoided having contact with [insert identified person]? What have you been worried about happening if you were in contact?
Have you got any concerns about the relationship? Have you got the sort of relationship you would like to have? How would you like things to be different? Have you got any idea how you could change things to have the sort of relationship you would like?
Are your differences so great that you might end the relationship?
[Note: if the relationship has ended, the problem area is still 'difficulties in relationship with x' but the status of the relationship is different. For example, if the parties are no longer in contact then the issue would be difficulties associated with not being in contact. These might be difficulties around feelings (are they missing them vs. feeling pleased that it was the right decision?), their self-esteem (do they feel rejected vs. more confident about self now not being undermined in the relationship), socialising (are they avoiding going out for fear of seeing them again vs. going out more now not have to worry about being jealous) etc.…You might need to say 'It sounds as if the relationship is over. Do you still think about the relationship even though it's over? Has it been on your mind at all over the last four weeks? Do you have any feelings about this even now? Do these cause you any problem? Do you wish things were different? Do you feel bad about not having contact with each other'?
[Note that differing role expectations need not be expressed in explicit/open 'conflict' but may reflect difficulties regarding communication and expectations.]

Interpersonal deficits (four weeks prior to treatment or previous four weeks)

[This interpersonal problem area includes anyone who appears to have a general lack of relationships with other people. It also includes anyone who does not feel close to people with whom relationships already exist or has concerns with regards to an intimate relationship e.g. wanting an intimate relationship or being able to keep it going. The word 'relationship' should be used interchangeably with the word 'friendship' or 'relationship with a partner' if it is deemed that this is more appropriate. Note that these questions continue to concern the four week time-frame rather than more generally, even if the problems have/have not been longstanding]
How much of the time do you spend with other people?
Who do you enjoy spending your time with?
How do you spend time with other people/friends? What do you enjoy doing with them?
How do you keep in contact with other people/friends? Email? Text? Phone call? Internet?
Do you enjoy the relationships that you have with people? Would you like them to be different in any way?
Do you ever feel lonely or isolated? How much of the time do you feel lonely or isolated?
Do you have people or friends that you feel close to?
Are you able to get close to other people? OR Did you feel close to people over these four weeks?
Do you have people you can confide in, who you can turn to when you need support or when you want to share a piece of news?
Is there anyone who confides in you?
How do you find meeting new people? Do you find it difficult?
Do you know what you want from a relationship? Have you an idea how you could achieve this?
Are you able to have a satisfying and meaningful intimate relationship?
Are you able to keep relationships going? Or do your relationships tend to break up fairly quickly?
Do you find yourself rapidly falling in and out of relationships?
Do you feel you are able to form relationships with people who provide you with what you want? Or do you keep finding yourself in unsuitable relationships?

Role transitions (four weeks prior to treatment or previous four weeks)
What sorts of things are you having difficulty adjusting to / are difficult now [change] has happened?
Is there anything else that is difficult about the change?
Does thinking about the change affect things?
Have you felt able to adjust to this change in your life?
Have you been able to get on with things you usually do since this change has taken place? Go to work? Meet with friends? Keep house, etc.
Have you been trying to avoid thinking about [the change] or things that remind you of it? Has this avoidance been difficult / caused you any difficulties?
Do you enjoy aspects of this change?
Can you see any good things about it?
Are you missing the way things used to be?
When did it happen?
Is it getting any easier?

Grief (four weeks prior to treatment or previous four weeks)
Are you spending time thinking about [insert the person concerned]?
How much time are you spending thinking about him/her?
When did [insert person identified in problem] die?
Have you been thinking about their death? What sorts of things have you been thinking about?
How have you been feeling as a result of this loss?
Do you miss them? / Were you missing them (over the four weeks)?
Have you been trying to avoid thinking about them? Are you avoiding anything that reminds you of them? Has avoiding thinking about them / things that remind you of them been difficult? Has it caused you any difficulties?
Are you able to think about the enjoyable parts of your relationship?
Are you able to think of the problems and difficulties in this relationship?
Are you involved in activities that you enjoyed before you lost [insert person identified in problem]?

Have you found ways to help fill the loss?
Have you been thinking about other losses in your life or been reminded of them?
Do you keep ……. belongings untouched (i.e., the same as before)?
Did you turn to anyone for support during this time? Were you able to talk to anyone about what happened?

Life goals (four weeks prior to treatment or previous four weeks)
I am going to be asking you about your 'life goals'. That is the goals you have for what you would like to be doing in your life. How much of a problem has it been that…
'…you are not doing what you would like to be doing in your life?' [and/or]
'…you do not know what you would like to be doing in your life?' [including: career, lifestyle, daily activities, interests]
Can you me more about what the problem is?
Do you know what you would like to be doing in your life? Do you feel you have a sense of what you are aiming for in life? Of where you are going?
Do you feel as if you are 'lost at sea' in terms of what you are doing in your life and where you are heading?
What problems does this it cause you?
Are you currently doing what you would like to be doing?
Do you feel that your current daily activities – work or study or role in life – are in accordance with your own values and what is important to you?
Do you feel interested in your current daily activities? Are they of value to you and important? Do they keep you busy and focused? …give you a sense of purpose?
Do you feel that you are making the most of your talents and skills?
What problems does it cause you if you are not doing what you would like to be doing?
Do you know how to work towards your life goals?
Are you already heading towards meeting your life goals?
At the moment are you able to make the most of opportunities to work towards your life goals (i.e., opportunities that crop up which might help you get to where you would like to be)?
Do you find your life goals are often changing? Does this make it difficult for you?
Do you know where to go, or who to ask for help, in meeting and developing your life goals? Do you feel there are barriers in the way of you achieving your life goals?

2. INTERPERSONAL SELF-EFFICACY IN RELATION TO THE IDENTIFIED PROBLEM
[This item concerns the individual's self-efficacy concerning resolving the identified interpersonal problem. Note that this is different from the severity of the problem and the participant's success or otherwise at overcoming it, although the two may be linked. Question 2 should be asked for each interpersonal problem.]

[If the participant believes that there is no longer a problem at all (which would have resulted in a rating of 0 above) skip 2a and ask 2b instead. For all other participants, ask 2a only.]

<u>Thinking about the past four weeks / four weeks before treatment began, to what extent have you believed that you are capable of overcoming or solving [insert problem] within the next year or two?</u>
Was this generally your belief over the last four weeks?

[If the participant appears not to understand what is being asked, the following clarificatory questions may be asked]:
If the participant struggles to understand that we are rating confidence, probe with:
I am interested in finding out how *confident* you are about overcoming or solving [insert problem] within the next year or two. I am asking you about your level of confidence over the last four weeks/four weeks before treatment began.
If the participant answers in terms of how much they have succeeded in overcoming it, then ask:

I am interested in the extent to which you *think* you <u>can</u> sort it out. This is not about whether you have actually been able to overcome ……. in last four weeks but instead about your confidence in your ability to overcome the problem.

If the participant says that his or her confidence is contingent on something else happening then he or she will need help to take this into account:
How likely do you think [insert the contingency] is to happen? What would need to take place for [insert the contingency] to happen? And *given this / taking this in to account*, to what extent have you believed (over this last four weeks) this you could overcome or solve [insert problem] within the next year or two?

If the participant asks you to explain what 'overcoming' means, then probe with:
By 'overcoming' I mean resolving the problem in a way that you feel would mean it had been dealt with successfully. It does not have to mean that [insert problem] has totally gone but that you feel it has been addressed in such a way that it is no longer a problem.

If the participant explains that the problem has not been present in the last four weeks because of circumstances (e.g. they are no longer living with the other party) then probe with:
Even if the situation has been different in the last four weeks, could you still tell me how much you have believed you have been capable of overcoming or solving [insert problem] within the next year or two?
2a(i) [ ]
2a(ii) [ ]
2a(iii) [ ]
[Ask 2b ONLY IF the participant says that the problem has already been resolved and would therefore not be able to answer 2a. It is another way of establishing self-efficacy. Individuals who believe that the resolution of the problem was due to them, perhaps with the support of treatment, should be given a higher rating on this item.]

<u>2b. If you have already solved/overcome/made great progress in overcoming [insert problem], how much do you think that this has been due to changes you have made, perhaps with the help of your therapist?</u>

2b(i) Primary problem [ ]
2b(ii) Secondary problem [ ]
2b(iii) Tertiary Problem [ ]

Rating Scale for Interpersonal Self-Efficacy in Overcoming the Identified Problem & General Interpersonal Self-Efficacy

<u>Rating scheme for questions 2 & 3</u>

0 – Not at all - No confidence in ability to overcome the problem
1 - Very little confidence in ability to overcome the problem.
2 – Slight - A degree of confidence in ability to overcome the problem.
3 - Some confidence in ability to overcome the problem.
4 – Moderate - Reasonable confidence in ability to overcome the problem.
5 – Confident about overcoming the problem.
6 – Marked - Definite confidence in ability to overcome the problem.

3. GENERAL INTERPERSONAL SELF-EFFICACY
[This item concerns self-efficacy with respect to interpersonal problems in general.]

The next question is a more general question with regards to your personal life and relationships. However, given that [insert problem] is an aspect of your life and relationships, please take this into account alongside other aspects of your personal life and relationships.

Over the past four weeks, to what extent have you believed that you are capable of overcoming difficulties in general in your relationships and personal life? Examples might include changes in your life circumstances or difficulties in your relationships with other people … such as people you are close to, or work colleagues or family members?
Was this generally your belief over the last four weeks?

If a participant says that it would depend on the circumstance then probe again with the same examples given at the end of the question. This helps to get participants to think about their confidence in overcoming interpersonal problems of a similar severity. However, avoid giving a description of a specific life event since the rating concerns the participant's confidence in overcoming the sorts of difficulties that they envisage in their personal life.

If the participant talks about a range of areas of their life without giving a description of their general interpersonal self-efficacy, then ask:
Can you tell me about how much you have believed that you are capable of overcoming difficulties in your life and relationships when taking all these different areas in to account?

If the participant says that his or her confidence is contingent on something else happening, help them take this in to account by probing with:
How likely do you think [insert the contingency] is to happen? What would need to take place for [insert the contingency] to happen? And *given this / taking this in to account*, to what extent have you believed that you are able to overcome or solve [insert problem/s] within the next year or two?

If the participant starts to tell you about an approach to dealing with problems (e.g. avoiding them) without reference to the original question repeat the original question. Then add,
Is this approach to do with your level of confidence?

Use the same rating scheme given for question 2.

3 [ ]

Appendix 9: Interpersonal Problem Form

INTERPERSONAL PROBLEM FORM

Date: [ ][ ]-[ ][ ]-[ ][ ]

Instructions for therapist

- To complete before session 5
- **Remind the patient NOT TO TELL THE RESEARCHER ABOUT THEIR TREATMENT.**
- You are trying to identify the patient's most significant current interpersonal problem. This is something in their life and relationships which affects their eating, their mood and / or their self-esteem. We are particularly interested in an interpersonal problem which affects the eating problem but also want to make sure to identify most significant current interpersonal problem – which may not affect the eating directly but may do so via mood or low self-esteem.
- Need to check that this has been going on for more than only the last four weeks – for example it should not be something that will only crop up for a limited time or is infrequent and just happens to be present at the moment.

In IPT

If your patient is in IPT please complete this using the problem that has been agreed for treatment. Please check that the problem area has been clarified with your patient before their interpersonal interview.

Please give the following rationale before the initial interview:
"We are interested in finding out how the treatments we are offering work. We need to give the researcher a brief description of the problem we are working on in treatment. Over time we might decide to work on other things in treatment but for the purposes of research you will be asked about this problem. This interview is for research purposes only so don't worry about it or the standard questions you are asked as it is separate from your treatment. The researcher doesn't know whether or not we are talking about this issue in our treatment"

At the end of treatment please explain:
"As you know we are interested in finding out how the treatments we are offering work. The research interview will be about the same area(s) that you were asked questions about before – even if there have been many changes in this area or we have been working on something somewhat different during treatment. This interview is for research purposes only so don't worry about it or the standard questions you are asked as it is separate from your treatment. The researcher doesn't know whether or not we have been talking about this issue in our treatment."

In CBT

If your patient is in CBT you will need to ask the questions below in the form of a 5-10mins semi-structured interview to establish an interpersonal problem. PLEASE RECORD THIS INTERVIEW.

Please give the following rationale for this:
"We are interested in finding out how the treatments we are offering work. For this reason we need to ask everybody in treatment about important issues in their life and relationships. The areas you will be asked about are limited to specific research categories about life and relationships. We have to pick a category together which most closely corresponds to a difficulty in your life and relationships. This interview is for research purposes only so don't worry about it or the standard questions you are asked – even if you are asked questions about something that doesn't seem much of a problem to you. The researcher doesn't know whether or not we are talking about this problem in our treatment".

At the end of treatment please explain:

"As you know we are interested in finding out how the treatments we are offering work. The research interview will be about the same area(s) that you were asked questions about before – even if there have been many changes in this area. This interview is for research purposes only so don't worry about it or the standard questions you are asked as it is separate from your treatment. The researcher doesn't know whether or not we have been talking about this issue in our treatment."

Questions

- ***Can you think of some aspect of your life or relationships that currently affect you and your eating? (focusing especially on the past four weeks)***
  For example…
    - *Having arguments or being in conflict with someone (e.g., with a friend, your partner, someone in your family, a colleague)*
    - *Missing someone who has left your life*
    - *Feeling lonely and isolated from other people*
    - *Having difficulty adjusting to a change in your life (e.g., leaving home, getting married, moving to a new area, taking up a new job)?*
    - *Feeling unsure about your goals in life or feeling as if you aren't doing something which is line with your goals*

  If there is more than one example ….
- ***Of the things that you have mentioned, which has the most influence on you and your eating?***

If there seems to be a clear primary problem area then use this. However, if you are unsure which of two problem areas may be most important then use both of these.

Provide below a brief description of the problem using patient's own words (e.g. Problems in my marriage; Missing my best friend; Being lonely; Getting used to being away from home, Having a stable intimate relationship with a man) and also indicate the key interpersonal problem(s) this most closely corresponds to (conflict, role transition, grief, deficit - lack of relationships, deficit – lack of stable relationships, life goals).

……
……
……

When completed please hand this to the mediational researcher – DO NOT TELL THE RESEARCHER WHICH TREATMENT THE PATIENT IS HAVING.

Appendix 10: Intervention Record Sheet for Enhanced Cognitive Behaviour Therapy

CBT-E Intervention Record Sheet

Patient initials _ _ Session No. _ _ Date _ _ /_ _ / _ _ Weight (BMI) _ _ _ (_ _ _ _)
Trial No. _ _ _ Week No. _ _ Session freq _ _ _ _ Previous weight _ _ _ (_ _ _ _)

Stage of treatment 1□ 2□ 3□ 4□ FU□ 20 week CBT / 40 week CBT F / B (circle) Recording codes:

Plan for the session

1.	6.
2.	7.
3.	8.
4.	9.
5.	10.

Session Content	Review
	Points for supervision Clinical Action points e.g. 10 week Rx progress letter □ End of Rx letter □ notes □ Other…… Medication (drug & dose): Research Action points Receipt of monitoring records □ Weekly q'aire □ IPA given to RA (by #4) Other……

Next Steps

Points to note for next session

Interventions used in this session

Weekly Weighing (by #1) □	Review of Progress □	Feeling Fat □
Regular Eating (+not in gaps) □	(Under-Eating) □	Analysis of Eating Changes □
Dietary Rules (avoided foods) □	(Weight Regain/Maintenance □	Mood Modulation/Tolerance □
Shape Checking: Monitoring □ Avoidance □ Comparisons □ □ Addressing		Maintaining Changes □
(Motivation to Change) □	Self-Evaluation Pie Chart □	Problem Solving □
Significant Other Session □	Marginalised Domains □	Historical Review □

Appendix 11: Intervention Record Sheet for Interpersonal Psychotherapy for Eating Disorders

IPT-ED Intervention Record Sheet

Patient initials _ _	Session No. _ _	Date _ _ /_ _ / _ _	Treatment phase 1□ 2□ 3□ FU□
Trial No. _ _ _	Tx week _ _ Elapsed week_ _	Session freq _____	Rec. codes_____

Problem Area:

Key issues:

Plan for the session

Session Content

Points for supervision

Points to note for next session

Clinical Action points
e.g. □ 10 week Rx progress letter □ End of Rx letter □ notes Other……

Medication (drug & dose):

Research Action points
Receipt of monitoring records □ Weekly q'aire □ IPA given to RA (by #4)
Other……

Appendix 12: Monitoring Record for Enhanced Cognitive Behaviour Therapy

CBT-E Monitoring Record

Instructions

The purpose of self-monitoring is two-fold:

1. It provides a detailed picture of how you eat, thereby bringing to your attention and that of your therapist the exact nature of your eating problem
2. By making you aware of what you are doing at the very time that you are doing it, self-monitoring helps you change behaviour that previously seemed outside your control.

Accurate self-monitoring is most important if these goals are to be met. At first, writing down everything down may be irritating and inconvenient, but it will soon become easier and of obvious value. We have yet to encounter anyone whose lifestyle makes monitoring truly impossible.

A sample monitoring sheet is shown overleaf. You may need to fill out two or more sheets per day. A new sheet should be started each day.

- Column 1 is for noting the time that you eat or drink anything.
- Column 2 is for recording the nature of the food and liquid consumed. You should provide a simple (non-technical) description of what you ate. Calories should not be recorded. Each item should be written down as soon as possible after its consumption. Recalling what you ate or drank some hours earlier is not satisfactory, since it will not help you change your behaviour. **Episodes of eating that you view as meals should be identified with brackets.** Snacks and other episodes of eating should not be bracketed.
- Asterisks should be placed in Column 3 adjacent to any episodes of eating that you felt at the time were excessive. It is essential to record all the food that you eat during binges.
- Column 4 should specify where the food or liquid was consumed. If this was in your home, the room should be specified.
- Column 5 is for recording when you vomit or take laxatives or diuretics (water tablets).
- Column 6 should be used as a diary to record events that influenced your eating: for example, if an argument precipitated a binge, you should note that down. You may wish to record other important events, even if they had no effect on your eating. Indeed, it is a good idea noting down all strong feelings; for example, feelings of depression, anxiety, boredom or loneliness or feeling fat. Concerns about shape or weight should also be noted here. Column 6 should also be used to record your weight each time you weigh yourself.

Obviously, if you are to record in the way that you are being asked, you will need to carry your monitoring sheets with you.

Every treatment session will include a review of your latest monitoring sheets (up to the time of the appointment). You should therefore remember to bring all of them with you.

Example of Monitoring Sheet

Date _ _ / _ _ / _ _ Day……

Time	Food and drink consumed	*	V/L	Place	Context and Comments

Appendix 13: Monitoring Record Interpersonal Psychotherapy for Eating Disorders

IPT-ED Monitoring Record

Instructions

The purpose of self-monitoring is three-fold:

1. It provides a detailed picture of what is happening in your life and relationships, thereby bringing to your attention (and that of your therapist) the exact nature of any relationship or life difficulties that you may be facing.
2. It makes you more aware of what is happening in your life and relationships at the very time that this is going on.
3. By recording your eating alongside events in your life and relationships, this will help you (and your therapist) make links between your eating and what is happening in your life and relationships. This will aid understanding of what aspects of your life and relationships may be contributing to keeping your eating problem going.

Accurate self-monitoring is most important if these goals are to be met. At first, writing everything down may be irritating and inconvenient, but it will soon become easier and of obvious value. We have yet to encounter anyone whose lifestyle makes monitoring truly impossible.

A sample monitoring sheet is shown overleaf. You may need to fill out two or more sheets per day. A new sheet should be started each day.

- Column 1 is for noting the time that you eat anything.
- Column 2 is for recording the nature of the food and liquid consumed. You should provide a simple (non-technical) description of what you ate. Calories should not be recorded. Each item should be written down as soon as possible after its consumption. Recalling what you ate or drank some hours earlier is not satisfactory. **Episodes of eating that you view as meals should be identified with brackets.** Snacks and other episodes of eating should not be bracketed.
- Column 3 - Asterisks should be placed in Column 3 adjacent to any episodes of eating that you felt at the time were excessive.
- Column 4 is for recording when you vomit or take laxatives or diuretics (water tablets).
- Column 5 should be used as a diary to record events in your life and relationships, especially those that have influenced your eating; for example, if an argument precipitated a binge or not eating, you should note that down. You should also record other important events in your life and relationships, even if they may have had no obvious effect on your eating. Indeed it is a good idea to note down any events which were associated with strong feelings. Strong feelings, such as feeling low, anxious, bored or lonely should also be recorded.

Obviously, if you are to record in the way that you are being asked, you will need to carry your monitoring sheets with you.

Every treatment session will include a review of your latest monitoring sheets (up to the time of the appointment). You should therefore remember to bring them with you.

Example of Monitoring Sheet

Date _ _ / _ _ / _ _ Day……

Time	Food and drink consumed	*	V/L	Relationships and Life

Statistical Output and Stata Code for Study 1: Regularity of Eating as a Mediator of the Effect of Enhanced Cognitive Behaviour Therapy on Frequency of Binge Eating

Appendix 14: Stata code for study 1a multilevel models and their coefficient plots

```
**********CODE FOR FINAL MULTILEVEL MODELS**********
/*
Notes on variables.
- trialno = patient ID
- dummy_icontrol = treatment allocation where 0 is IPT-ED, and 1 is CBT-E
- regulareating = regularity of eating
- center_regulareating = regularity of eating (centred around each patient's score at baseline)
- wklyOBE = frequency of binge eating
- txweek = actual treatment week
- dummy_txweek = treatment week, where zero indicates baseline (start of treatment)
- dummy_txweek2 = treatment week as a quadratic trend
- dummy_diag = eating disorder diagnosis (bulimia nervosa vs. eating disorder not otherwise
specified)
- mean_edewc = level of  weight concern at baseline centred at the sample mean
- everbinge = presence of binge eating in the month prior to treatment
- mean_age = age at baseline centred at the sample mean
- mean_bdi = depression at baseline centred at the sample mean
*/

****STEP 1- IMPACT OF THE INTERVENTION ON THE PURPORTED MEDIATOR****
*run final model
bootstrap, cluster(trialno) reps(1000) seed(686922911): xtmixed regulareating c.dummy_txweek
c.dummy_txweek2 dummy_icontrol c.dummy_txweek#dummy_icontrol
c.dummy_txweek2#dummy_icontrol dummy_diag mean_bdi || trialno: dummy_txweek if txweek<5 ,
reml cov(unstr) var
estat bootstrap, bc
estat icc
estat ic
*create summary table of model
eststo est1
estout est1, cells("b(fmt(2)) se p(fmt(3)) ci_bc[1] ci_bc[2]") modelwidth(9) style(fixed) stats(aic
N_clust N N_reps)
*create coefficient plot of key variables
coefplot est1, omitted keep(dummy_txweek dummy_txweek2 1.dummy_icontrol#?.dummy_txweek
1.dummy_icontrol#?.dummy_txweek2) xline(0) ci(ci_bc) ciopts(recast(rcap)) mlabel format(%9.2f)
mlabposition(12) mlabgap(*1) name(REMLMcoeffinterval, replace) xtitle("{bf:Effect on Regularity of
Eating}", height(5) )title("{bf:1. Impact of Intervention on Purported Mediator}", size(medlarge)
color(black) span) coeflabel(dummy_txweek = "Treatment week"  dummy_txweek2 = "Treatment
week (quadratic trend)" 1.dummy_icontrol#?.dummy_txweek ="CBT-E (linear trend)", wrap(40))
order (dummy_txweek dummy_txweek2 1.dummy_ccontrol#?.dummy_txweek
1.dummy_icontrol#?.dummy_txweek2) headings(dummy_txweek  = "{bf:Impact of Treatment
Week}" 1.dummy_icontrol#?.dummy_txweek  = "{bf:Impact of Treatment Allocation}")
saving(REMLMcoeffinterval, replace)
```

```
****STEP 2 - IMPACT OF THE INTERVENTION ON THE OUTCOME****
*run final model
bootstrap, cluster(trialno) reps(1000) seed(686922911): xtmixed wklyOBE  c.dummy_txweek
c.dummy_txweek2 dummy_icontrol dummy_icontrol#c.dummy_txweek
dummy_icontrol#c.dummy_txweek2 c.mean_edewc dummy_diag dummy_diag#c.dummy_txweek
everbinge everbinge#c.dummy_txweek mean_age c.mean_age#c.dummy_txweek || trialno:
dummy_txweek if txweek<5, reml cov(unstr) var
estat bootstrap, bc
estat icc
estat ic
*create summary table of model
eststo est2
estout est2, cells("b(fmt(2)) se p(fmt(3)) ci_bc[1] ci_bc[2]") modelwidth(9) style(fixed) stats(aic
N_clust N N_reps)
*create coefficient plot of key variables
coefplot est2, omitted keep(dummy_txweek dummy_txweek2 1.dummy_icontrol#c.dummy_txweek
1.dummy_icontrol#c.dummy_txweek2) xline(0) ci(ci_bc) ciopts(recast(rcap)) mlabel format(%9.2f)
mlabposition(12) mlabgap(*1) name(OBEMLMcoeffinterval, replace) xtitle("{bf:Effect on Frequency
of Binge Eating}", height(5)) title("{bf:2. Impact of Intervention on Outcome}", size(medlarge)
color(black) span) coeflabel(dummy_txweek = "Treatment week" dummy_txweek2 = "Treatment
week (quadratic trend)" 1.dummy_icontrol#?.dummy_txweek = "CBT-E (linear trend)"
1.dummy_icontrol#?.dummy_txweek2 = "CBT-E (quadratic trend)", wrap(40)) order (dummy_txweek
dummy_txweek2 1.dummy_icontrol#?.dummy_txweek 1.dummy_icontrol#?.dummy_txweek2)
headings(dummy_txweek  = "{bf:Impact of Treatment Week}" 1.dummy_icontrol#?.dummy_txweek
= "{bf:Impact of Treatment Allocation}") saving(OBEMLMcoeffinterval, replace)

****STEP 3 - IMPACT OF THE PURPORTED MEDIATOR ON THE OUTCOME****
*run final model
bootstrap, cluster(trialno) reps(1000) seed(686922911): xtmixed wklyOBE  c.dummy_txweek
center_regulareating c.mean_edewc dummy_diag dummy_diag#c.dummy_txweek everbinge
mean_age c.mean_age#c.dummy_txweek || trialno: dummy_txweek if txweek<5, reml cov(unstr) var
estat bootstrap, bc
estat icc
estat ic
*create summary table of model
eststo est3
estout est3, cells("b(fmt(2)) se p(fmt(3)) ci_bc[1] ci_bc[2]") modelwidth(9) style(fixed) stats(aic
N_clust N N_reps)
*create coefficient plot of key variables
coefplot est3, omitted baselevels keep(dummy_txweek dummy_txweek2 center_regulareating)
xline(0) ci(ci_bc) ciopts(recast(rcap)) mlabel format(%9.2f) mlabposition(12) mlabgap(*1)
name(MOREMLMcoeffinterval, replace) xtitle("{bf:Effect on Frequency of Binge Eating}", height(5))
title("{bf:3. Impact of Purported Mediator on Outcome}", size(medlarge) color(black) span)
coeflabel(dummy_txweek = "Treatment week" dummy_txweek2 = "Treatment week (quadratic
trend)" center_regulareating = "Regularity of eating", wrap(40)) order (dummy_txweek
dummy_txweek2 center_regulareating) headings(dummy_txweek = "{bf:Impact of Treatment Week}"
center_regulareating = "{bf:Impact of Purported Mediator}") saving(REFWMLMcoeffinterval, replace)

****STEP 4 - IMPACT OF THE INTERVENTION ON THE OUTCOME WHEN INCLUDING THE PURPORTED
MEDIATOR IN THE MODEL (CONTEMPORANEOUS RELATIONSHIP)****
*run final model
bootstrap, cluster(trialno) reps(1000) seed(686922911): xtmixed wklyOBE  c.dummy_txweek
c.dummy_txweek2 dummy_icontrol dummy_icontrol#c.dummy_txweek
dummy_icontrol#c.dummy_txweek2 center_regulareating c.mean_edewc dummy_diag
```

```
dummy_diag#c.dummy_txweek everbinge everbinge#c.dummy_txweek mean_age
c.mean_age#c.dummy_txweek || trialno: dummy_txweek if txweek<5, reml cov(unstr) var
estat bootstrap, bc
estat icc
estat ic
*create summary table of model
eststo est4
estout est4, cells("b(fmt(2)) se p(fmt(3)) ci_bc[1] ci_bc[2]") modelwidth(9) style(fixed) stats(aic
N_clust N N_reps)
 *create coefficient plot of key variables
coefplot est4, omitted baselevels keep(dummy_txweek dummy_txweek2
1.dummy_icontrol#c.dummy_txweek 1.dummy_icontrol#c.dummy_txweek2 center_regulareating)
xline(0) ci(ci_bc) ciopts(recast(rcap)) mlabel format(%9.2f) mlabposition(12) mlabgap(*1)
name(medREMLMcoeffinterval, replace) xtitle("{bf:Effect on Frequency of Binge Eating}", height(5))
title("{bf:4. Impact of Intervention on Outcome with Purported Mediator}", size(medlarge)
color(black) span) coeflabel(dummy_txweek = "Treatment week"  dummy_txweek2 = "Treatment
week (quadratic trend)" 1.dummy_icontrol#?.dummy_txweek = "CBT-E (linear trend)"
1.dummy_icontrol#?.dummy_txweek2 = "CBT-E (quadratic trend)" center_adj_eqq16 = "Frequency of
weighing" center_regulareating = "Regularity of eating", wrap(40)) order (dummy_txweek
dummy_txweek2 1.dummy_icontrol#?.dummy_txweek 1.dummy_icontrol#?.dummy_txweek2
center_regulareating) headings(dummy_txweek  = "{bf:Impact of Treatment
Week}"1.dummy_icontrol#?.dummy_txweek  = "{bf:Impact of Treatment Allocation}"
center_regulareating = "{bf:Impact of Purported Mediator}") saving(medREMLMcoeffinterval,
replace)

****COMBINE COEFPLOTS INTO ONE GRAPH****
graph combine REMLMcoeffinterval OBEMLMcoeffinterval MOREMLMcoeffinterval
medREMLMcoeffinterval, name(combRE_MLM, replace) xsize(20) ysize(10) saving(combRE_MLM,
replace)
```

Appendix 15: Output for study 1a, step 1. Multi-level model of the impact of the intervention (CBT-E) on the purported mediator (regularity of eating)

Model Components	Variable Details	B	Boot. SE	p	Bias-corrected 95% CI of B	
					Lower	**Upper**
Baseline predictors	Treatment allocation (CBT-E vs. IPT-ED)	0.71	0.19	.000	0.34	1.07
	Eating disorder diagnosis (BN vs. EDNOS)	-0.59	0.16	.000	-0.92	-0.27
	Depression	-0.02	0.01	.019	-0.04	-0.00
Impact of treatment week	Treatment week (linear trend)	0.19	0.13	.150	-0.06	0.45
	Treatment week (quadratic trend)	-0.06	0.04	.197	-0.14	0.02
Impact of treatment allocation	CBT-E*treatment week (linear trend)	1.12	0.21	.000	0.68	1.51
	CBT-E*treatment week (quadratic trend)	-0.23	0.07	.001	-0.34	-0.09
Constant	Patients in IPT with EDNOS average depression at baseline	2.76	0.15	.000	2.45	3.04
Random effects (log values)	SD (treatment week)	-1.78	0.11	.000	-1.91	-1.91
	SD (constant)	0.05	0.06	.411	-0.12	0.12
	Correlation (treatment week*constant)	-0.19	0.10	.043	-0.27	0.05
	SD (residual)	-0.50	0.06	.000	.	.
Statistics	AIC	1175.14				
	Number of patients	117				
	Number of observations	439				
	Number of bootstraps	1000/1000				

Notes. B=unstandardized coefficient, *p*-value based on 95% confidence interval, CI=confidence interval. BN=bulimia nervosa. EDNOS=eating disorder not otherwise specified.

Appendix 16: Output for study 1a, step 2. Multi-level model of the impact of the intervention (CBT-E) on the outcome (frequency of binge eating)

Model Components	Variable Details	B	Boot. SE	p	Bias-corrected 95% CI of B	
					Lower	**Upper**
Baseline predictors	Treatment allocation (CBT-E vs. IPT-ED)	-0.43	0.43	.316	-1.31	0.41
	Weight concern	-0.41	0.13	.002	-0.66	-0.14
	Eating disorder diagnosis (BN vs. EDNOS)	2.43	0.49	.000	1.51	3.44
	Presence of binge eating	1.66	0.38	.000	0.97	2.47
	Age	-0.06	0.02	.009	-0.12	-0.02
Impact of treatment week	Treatment week (linear trend)	0.75	0.35	.032	0.12	1.50
	Treatment week (quadratic trend)	-0.25	0.12	.035	-0.50	-0.03
Impact of baseline predictors during treatment	BN*treatment week (linear trend)	-0.34	0.17	.048	-0.71	-0.02
	Presence of binge eating*treatment week (linear trend)	-0.20	0.10	.041	-0.40	-0.01
	Age*treatment week (linear trend)	0.02	0.01	.002	0.01	0.03
Impact of treatment allocation	CBT-E*treatment week (linear trend)	-1.40	0.51	.007	-2.47	-0.47
	CBT-E*treatment week (quadratic trend)	0.43	0.16	.006	0.12	0.75
Constant	Patients in IPT-ED with EDNOS, no binge eating, average weight concern and age at baseline	0.56	0.34	.105	-0.19	1.16
Random effects (log values)	SD (treatment week)	-0.47	0.11	.000	-0.73	-0.47
	SD (constant)	0.88	0.08	.000	0.61	0.97
	Correlation (treatment week*constant)	-0.63	0.13	.000	-0.86	-0.32
	SD (residual)	0.50	0.09	.000	0.51	0.57
Statistics	AIC	2060.64				
	Number of patients	119				
	Number of observations	450				
	Number of bootstraps	1000/1000				

Notes. B=unstandardized coefficient, *p*-value based on 95% confidence interval, CI=confidence interval. BN=bulimia nervosa. EDNOS=eating disorder not otherwise specified.

Appendix 17: Output for study 1a, step 3. Multi-level model of impact of the purported mediator (regularity of eating) on the outcome (frequency of binge eating)

Model Components	Variable Details	B	Boot. SE	p	Bias-corrected 95% CI of B	
					Lower	**Upper**
Baseline predictors	Weight concern	-0.38	0.14	.008	-0.65	-0.10
	Eating disorder diagnosis (BN vs. EDNOS)	2.54	0.49	.000	1.60	3.50
	Presence of binge eating	1.33	0.30	.000	0.70	1.90
	Age	-0.06	0.03	.021	-0.11	-0.01
Impact of treatment week	Treatment week (linear trend)	-0.03	0.08	.709	-0.18	0.12
Impact of baseline predictors during treatment	BN*treatment week (linear trend)	-0.45	0.15	.003	-0.76	-0.17
	Age*treatment week (linear trend)	0.02	0.01	.001	0.01	0.03
Impact of purported mediator	Regularity of eating	-0.57	0.13	.000	-0.86	-0.32
Constant	Patients with EDNOS, no binge eating, average weight concern and age at baseline	0.72	0.27	.006	0.25	1.29
Random effects (log values)	SD (treatment week)	-0.51	0.11	.000	-0.88	-0.51
	SD (constant)	0.90	0.08	.000	0.65	0.99
	Correlation (treatment week*constant)	-0.64	0.12	.000	-0.88	-0.38
	SD (residual)	0.48	0.10	.000	0.47	0.60
Model statistics	AIC	1998.40				
	Number of patients	116				
	Number of observations	441				
	Number of bootstraps	1000/1000				

Notes. B=unstandardized coefficient, *p*-value based on 95% confidence interval, CI=confidence interval. BN=bulimia nervosa. EDNOS=eating disorder not otherwise specified.

Appendix 18: Output for study 1a, step 4. Multi-level model of impact of the intervention (CBT-E) on the outcome (frequency of binge eating) when accounting for the purported mediator (regularity of eating)

Model Components	Variable Details	B	Boot. SE	p	Bias-corrected 95% CI of B	
					Lower	**Upper**
Baseline predictors	Treatment allocation (CBT-E vs. IPT-ED)	-0.49	0.43	.264	-1.36	0.30
	Weight concern	-0.37	0.14	.009	-0.64	-0.08
	Eating disorder diagnosis (BN vs. EDNOS)	2.38	0.49	.000	1.42	3.33
	Presence of binge eating	1.73	0.38	.000	1.08	2.55
	Age	-0.06	0.03	.012	-0.12	-0.01
Impact of treatment week	Treatment week (linear trend)	0.79	0.34	.021	0.17	1.51
	Treatment week (quadratic trend)	-0.26	0.12	.024	-0.52	-0.06
Impact of baseline predictors during treatment	BN*treatment week (linear trend)	-0.39	0.17	.020	-0.73	-0.08
	Presence of binge eating*treatment week (linear trend)	-0.17	0.11	.118	-0.39	0.03
	Age*treatment week (linear trend)	0.02	0.01	.001	0.01	0.03
Impact of treatment allocation	CBT-E*treatment week (linear trend)	-0.71	0.52	.170	-1.71	0.29
	CBT-E*treatment week (quadratic trend)	0.29	0.16	.067	-0.01	0.60
Impact of purported mediator	Regularity of eating	-0.56	0.15	.000	-0.87	-0.29
Constant	Patients in IPT-ED with EDNOS, no of binge eating, average weight concern and age at baseline	0.59	0.34	.081	-0.08	1.28
Random effects (log values)	SD (treatment week)	-0.49	0.11	.000	-0.83	-0.49
	SD (constant)	0.90	0.08	.000	0.65	0.99
	Correlation (treatment week*constant)	-0.63	0.12	.000	-0.87	-0.37
	SD (residual)	0.47	0.09	.000	0.47	0.56
Statistics	AIC	2007.53				
	Number of patients	116				
	Number of observations	441				
	Number of bootstraps	1000/1000				

Notes. B=unstandardized coefficient, *p*-value based on 95% confidence interval, CI=confidence interval. BN=bulimia nervosa. EDNOS=eating disorder not otherwise specified.

Appendix 19: Stata code for study 1b autoregressive structural equation model

```
*******CODE FOR FINAL AUTOREGRESSIVE STRUCTURAL EQUATION MODEL*******
/*
Notes on variables.
- trialno = patient ID
- dummy_icontrol = treatment allocation where 0 is IPT-ED, and 1 is CBT-E
- regulareatingwk1 = regularity of eating at week 1
- regulareatingwk2 = regularity of eating at week 2
- regulareatingwk3 = regularity of eating at week 3
- regulareatingwk4 = regularity of eating at week 4
- wklyOBEwk1 = frequency of binge eating at week 1
- wklyOBEwk2 = frequency of binge eating at week 2
- wklyOBEwk3 = frequency of binge eating at week 3
- wklyOBEwk4 = frequency of binge eating at week 4
- dummy_diag = eating disorder diagnosis (bulimia nervosa vs. eating disorder not otherwise
specified)
- mean_edewc = level of  weight concern at baseline centred at the sample mean
- everbinge = presence of binge eating in the month prior to treatment
- mean_age = age at baseline centred at the sample mean
- mean_bdi = depression at baseline centred at the sample mean
*/
*unstandardised
sem (regulareatingwk1->regulareatingwk2 wklyOBEwk2) (regulareatingwk2->regulareatingwk3
wklyOBEwk2 wklyOBEwk3) (regulareatingwk3->regulareatingwk4 wklyOBEwk3 wklyOBEwk4)
(regulareatingwk4->wklyOBEwk4) (wklyOBEwk1->wklyOBEwk2) (wklyOBEwk2->wklyOBEwk3)
(wklyOBEwk3->wklyOBEwk4) (dummy_icontrol mean_bdi dummy_diag->regulareatingwk2
regulareatingwk3 regulareatingwk4) (dummy_icontrol everbinge dummy_diag mean_edewc
mean_age->wklyOBEwk2 wklyOBEwk3 wklyOBEwk4), cov(e.regulareatingwk2*e.regulareatingwk3)
cov(e.regulareatingwk3*e.regulareatingwk4) cov(dummy_icontrol*regulareatingwk1)
cov(e.wklyOBEwk2*e.wklyOBEwk3) cov(e.wklyOBEwk3*e.wklyOBEwk4)
cov(dummy_icontrol*wklyOBEwk1) vce(bootstrap, reps(1000) seed(1242741187)) method(mlmv)
estat bootstrap, bc
*fit indices
estat gof, stats(all)
*test stability of path coefficients for regularity of eating
*RE
test(_b[regulareatingwk2:regulareatingwk1]==_b[regulareatingwk3:regulareatingwk2]==_b[regularea
tingwk4:regulareatingwk3])
*test stability of path coefficients for frequency of binge eating
test( _b[wklyOBEwk2:wklyOBEwk1]== _b[wklyOBEwk3:wklyOBEwk2]==
_b[wklyOBEwk4:wklyOBEwk3])
*get table of direct, indirect and total effects of treatment, set matrices & return scalars for model to
run bootstrap teffects for effect of CBT-E
estat teffects, standardized
mat list r(indirect)
mat list r(total)
program indireff, rclass
sem (regulareatingwk1->regulareatingwk2 wklyOBEwk2) (regulareatingwk2->regulareatingwk3
wklyOBEwk2 wklyOBEwk3) (regulareatingwk3->regulareatingwk4 wklyOBEwk3 wklyOBEwk4)
(regulareatingwk4->wklyOBEwk4) (wklyOBEwk1->wklyOBEwk2) (wklyOBEwk2->wklyOBEwk3)
(wklyOBEwk3->wklyOBEwk4) (dummy_icontrol mean_bdi dummy_diag->regulareatingwk2
regulareatingwk3 regulareatingwk4) (dummy_icontrol everbinge dummy_diag mean_edewc
mean_age->wklyOBEwk2 wklyOBEwk3 wklyOBEwk4), cov(e.regulareatingwk2*e.regulareatingwk3)
cov(e.regulareatingwk3*e.regulareatingwk4) cov(dummy_icontrol*regulareatingwk1)
```

```
cov(e.wklyOBEwk2*e.wklyOBEwk3) cov(e.wklyOBEwk3*e.wklyOBEwk4)
cov(dummy_icontrol*wklyOBEwk1) method(mlmv)
estat teffects
mat bi = r(indirect)
mat bt = r(total)
return scalar biwk2  = bi[1,8]
return scalar biwk3  = bi[1,24]
return scalar biwk4  = bi[1,43]
return scalar btwk2  = bt[1,8]
return scalar btwk3  = bt[1,24]
return scalar btwk4  = bt[1,43]
end
bootstrap r(biwk2) r(biwk3) r(biwk4) r(btwk2) r(btwk3) r(btwk4), reps(1000) seed(1242741187):
indireff
estat bootstrap, bc
program drop indireff
*standardised coeffs
sem (regulareatingwk1->regulareatingwk2 wklyOBEwk2) (regulareatingwk2->regulareatingwk3
wklyOBEwk2 wklyOBEwk3) (regulareatingwk3->regulareatingwk4 wklyOBEwk3 wklyOBEwk4)
(regulareatingwk4->wklyOBEwk4) (wklyOBEwk1->wklyOBEwk2) (wklyOBEwk2->wklyOBEwk3)
(wklyOBEwk3->wklyOBEwk4) (dummy_icontrol mean_bdi dummy_diag->regulareatingwk2
regulareatingwk3 regulareatingwk4) (dummy_icontrol everbinge dummy_diag mean_edewc
mean_age->wklyOBEwk2 wklyOBEwk3 wklyOBEwk4), cov(e.regulareatingwk2*e.regulareatingwk3)
cov(e.regulareatingwk3*e.regulareatingwk4) cov(dummy_icontrol*regulareatingwk1)
cov(e.wklyOBEwk2*e.wklyOBEwk3) cov(e.wklyOBEwk3*e.wklyOBEwk4)
cov(dummy_icontrol*wklyOBEwk1) vce(bootstrap, reps(1000) seed(1242741187)) method(mlmv)
standardized
```

Appendix 20: Output for study 1b. Autoregressive structural equation model of the relationships between the intervention (CBT-E), the purported mediator (regularity of eating; RE) and the outcome (frequency of binge eating; BE)

Model Components	B	Boot. SE	Bias-corrected 95% CI of B		β
			Lower	**Upper**	
RE1→RE2	0.64	0.08	0.48	0.79	0.57
TX→RE2	1.21	0.18	0.86	1.58	0.41
Depression→RE2	0.00	0.01	-0.02	0.02	0.02
ED diagnosis→RE2	-0.29	0.17	-0.67	0.03	-0.10
RE2→ BE2	-0.71	0.28	-1.26	-0.22	-0.33
RE1→BE2	0.31	0.25	-0.18	0.80	0.13
BE1→BE2	0.52	0.10	0.32	0.71	0.55
TX→ BE2	-0.11	0.53	-1.16	0.91	-0.02
ED diagnosis→ BE2	0.74	0.47	-0.22	1.74	0.11
BE Presence→ BE2	0.93	0.35	0.37	1.75	0.11
WC→ BE2	-0.10	0.18	-0.48	0.23	-0.04
Age→BE2	0.00	0.02	-0.04	0.05	0.00
RE2→RE3	1.01	0.08	0.86	1.20	0.92
TX→RE3	0.33	0.22	-0.12	0.70	0.10
Depression→RE3	-0.01	0.01	-0.02	0.01	-0.03
ED diagnosis→RE3	-0.18	0.17	-0.49	0.20	-0.05
RE2→ BE3	0.71	0.33	0.15	1.52	0.33
BE2→BE3	0.78	0.25	0.38	1.35	0.78
RE3→BE3	-0.76	0.28	-1.34	-0.23	-0.39
TX→BE3	-0.32	0.63	-1.88	0.75	-0.05
ED diagnosis→BE3	0.07	0.66	-1.20	1.38	0.01
BE Presence→BE3	-0.25	0.49	-1.35	0.61	-0.03
WC →BE3	0.03	0.19	-0.38	0.37	0.01
Age→BE3	-0.01	0.02	-0.05	0.03	-0.03
RE3→RE4	0.91	0.07	0.78	1.06	0.94
TX→RE4	0.25	0.24	-0.25	0.64	0.08
Depression→RE4	0.01	0.01	-0.01	0.02	0.06
ED diagnosis→RE4	0.17	0.15	-0.14	0.45	0.05
RE3→ BE4	0.53	0.28	-0.04	1.08	0.33
BE3→BE4	0.74	0.15	0.50	1.10	0.91
RE4→BE4	-0.71	0.31	-1.36	-0.18	-0.44
TX→BE4	0.95	0.48	0.07	1.92	0.18
ED diagnosis→BE4	-0.20	0.39	-0.96	0.62	-0.04
BE Presence→BE4	0.22	0.36	-0.54	0.91	0.03
WC →BE4	-0.28	0.18	-0.69	0.03	-0.15
Age→BE4	0.02	0.02	-0.01	0.07	0.07
Var(e.RE2)	0.79	0.09	0.67	1.04	0.36
Var(e.BE2)	4.19	0.64	3.25	6.05	0.41
Var(e.RE3)	0.62	0.11	0.46	0.92	0.23
Var(e.BE3)	4.43	1.32	2.63	7.71	0.44
Var(e.RE4)	0.52	0.10	0.36	0.74	0.21
Var(e.BE4)	3.49	1.06	1.96	7.60	0.52
Cov(e.RE2*e.RE3)	-0.32	0.10	-0.54	-0.14	-0.45

Model Components	B	Boot. SE	Bias-corrected 95% CI of B		β
			Lower	**Upper**	
Cov(e.BE2*e.BE3)	-0.82	0.93	-3.22	0.49	-0.19
Cov(e.RE3*e.RE4)	-0.21	0.08	-0.37	-0.08	-0.37
Cov(e.BE3*e.BE4)	-1.73	0.87	-4.16	-0.46	-0.44
Cov(RE1*BE1)	-1.79	0.43	-2.71	-0.93	-0.40
Cov(RE1*TX)	0.18	0.06	0.07	0.30	0.27
Cov(RE1*Depression)	-3.72	1.31	-6.37	-1.27	-0.26
Cov(RE1*ED diagnosis)	-0.13	0.06	-0.24	-0.02	-0.20
Cov(RE1*BE Presence)	-0.04	0.05	-0.12	0.07	-0.07
Cov(RE1*WC)	0.14	0.18	-0.17	0.57	0.08
Cov(RE1*Age)	0.10	1.02	-2.09	1.93	0.01
Cov(BE1*TX)	-0.12	0.15	-0.44	0.14	-0.07
Cov(BE1*Depression)	0.39	3.58	-6.71	7.14	0.01
Cov(BE1*ED diagnosis)	0.71	0.14	0.44	1.02	0.43
Cov(BE1*BE Presence)	0.43	0.09	0.26	0.64	0.33
Cov(BE1*WC)	-1.09	0.45	-2.00	-0.24	-0.23
Cov(BE1*Age)	-5.07	2.11	-9.34	-0.97	-0.20
Cov(TX*Depression)	-0.36	0.49	-1.42	0.49	-0.07
Cov(TX*ED diagnosis)	-0.01	0.02	-0.05	0.03	-0.05
Cov(TX*BE Presence)	0.01	0.02	-0.02	0.04	0.06
Cov(TX*WC)	0.07	0.06	-0.05	0.20	0.11
Cov(TX*Age)	-0.48	0.31	-1.11	0.13	-0.13
Cov(Depression*ED diagnosis)	0.59	0.46	-0.34	1.45	0.11
Cov(Depression*BE Presence)	0.50	0.35	-0.18	1.21	0.12
Cov(Depression*WC)	4.74	1.48	1.99	7.90	0.31
Cov(Depression*Age)	12.98	8.02	-2.77	28.76	0.16
Cov(ED diagnosis*BE Presence)	0.08	0.01	0.06	0.11	0.40
Cov(ED diagnosis*WC)	-0.02	0.06	-0.15	0.08	-0.03
Cov(ED diagnosis*Age)	-0.37	0.32	-0.98	0.25	-0.10
Cov(BE Presence*WC)	0.01	0.05	-0.08	0.09	0.01
Cov(BE Presence*Age)	-0.27	0.30	-0.93	0.26	-0.09
Cov(WC*Age)	0.45	1.10	-2.43	2.17	0.04
Successful bootstraps	892/1000				

Notes. B=unstandardized coefficient. SE=standard error. CI=confidence interval. β=standardised coefficient. RE# = regularity of eating where # indicates intervention week. BE# = frequency of binge eating where # indicates intervention week. TX=treatment allocation, BE Presence=Presence of binge eating in the month prior to starting treatment, WC = weight concern at baseline.

Appendix 21: Stata code for study 1c alternative autoregressive structural equation model

```
*******CODE FOR ALTERNATIVE AUTOREGRESSIVE STRUCTURAL EQUATION MODEL*******
/*
Notes on variables.
- trialno = patient ID
- dummy_icontrol = treatment allocation where 0 is IPT-ED, and 1 is CBT-E
- regulareatingwk1 = regularity of eating at week 1
- regulareatingwk2 = regularity of eating at week 2
- regulareatingwk3 = regularity of eating at week 3
- regulareatingwk4 = regularity of eating at week 4
- wklyOBEwk1 = frequency of binge eating at week 1
- wklyOBEwk2 = frequency of binge eating at week 2
- wklyOBEwk3 = frequency of binge eating at week 3
- wklyOBEwk4 = frequency of binge eating at week 4
- dummy_diag = eating disorder diagnosis (bulimia nervosa vs. eating disorder not otherwise
specified)
- mean_edewc = level of  weight concern at baseline centred at the sample mean
- everbinge = presence of binge eating in the month prior to treatment
- mean_age = age at baseline centred at the sample mean
- mean_bdi = depression at baseline centred at the sample mean
*/

*unstandardised
sem (regulareatingwk1->regulareatingwk2) (regulareatingwk2->regulareatingwk3) (regulareatingwk3-
>regulareatingwk4) (wklyOBEwk1->wklyOBEwk2 regulareatingwk2) (wklyOBEwk2->wklyOBEwk3
regulareatingwk2 regulareatingwk3) (wklyOBEwk3->wklyOBEwk4 regulareatingwk3
regulareatingwk4) (wklyOBEwk4->regulareatingwk4) (dummy_icontrol mean_bdi dummy_diag-
>regulareatingwk2 regulareatingwk3 regulareatingwk4) (dummy_icontrol everbinge dummy_diag
mean_edewc mean_age->wklyOBEwk2 wklyOBEwk3 wklyOBEwk4),
cov(e.regulareatingwk2*e.regulareatingwk3) cov(e.regulareatingwk3*e.regulareatingwk4)
cov(dummy_icontrol*regulareatingwk1) cov(e.wklyOBEwk2*e.wklyOBEwk3)
cov(e.wklyOBEwk3*e.wklyOBEwk4) cov(dummy_icontrol*wklyOBEwk1) method(mlmv)
vce(bootstrap, reps(1000) seed(1242741187))
estat bootstrap, bc
*fit indices
estat gof, stats(all)
*test stability of path coefficients for regularity of eating
test(_b[regulareatingwk2:regulareatingwk1]==_b[regulareatingwk3:regulareatingwk2]==_b[regularea
tingwk4:regulareatingwk3])
*test stability of path coefficients for frequency of binge eating
test( _b[wklyOBEwk2:wklyOBEwk1]== _b[wklyOBEwk3:wklyOBEwk2]==
_b[wklyOBEwk4:wklyOBEwk3])
*get table of direct, indirect and total effects of treatment, set matrices & return scalars for model to
run bootstrap teffects for effect of CBT-E
estat teffects, standardized
mat list r(indirect)
mat list r(total)
program indireff, rclass
sem (regulareatingwk1->regulareatingwk2) (regulareatingwk2->regulareatingwk3) (regulareatingwk3-
>regulareatingwk4) (wklyOBEwk1->wklyOBEwk2 regulareatingwk2) (wklyOBEwk2->wklyOBEwk3
regulareatingwk2 regulareatingwk3) (wklyOBEwk3->wklyOBEwk4 regulareatingwk3
regulareatingwk4) (wklyOBEwk4->regulareatingwk4) (dummy_icontrol mean_bdi dummy_diag-
>regulareatingwk2 regulareatingwk3 regulareatingwk4) (dummy_icontrol everbinge dummy_diag
mean_edewc mean_age->wklyOBEwk2 wklyOBEwk3 wklyOBEwk4),
```

```
cov(e.regulareatingwk2*e.regulareatingwk3) cov(e.regulareatingwk3*e.regulareatingwk4)
cov(dummy_icontrol*regulareatingwk1) cov(e.wklyOBEwk2*e.wklyOBEwk3)
cov(e.wklyOBEwk3*e.wklyOBEwk4) cov(dummy_icontrol*wklyOBEwk1) method(mlmv)
  estat teffects
mat bi = r(indirect)
mat bt = r(total)
return scalar biwk2  = bi[1,4]
return scalar biwk3  = bi[1,15]
return scalar biwk4  = bi[1,28]
return scalar btwk2  = bt[1,4]
return scalar btwk3  = bt[1,15]
return scalar btwk4  = bt[1,28]
end
bootstrap r(biwk2) r(biwk3) r(biwk4) r(btwk2) r(btwk3) r(btwk4), reps(1000) seed(1242741187):
indireff
estat bootstrap, bc
program drop indireff
*standardised
sem (regulareatingwk1->regulareatingwk2) (regulareatingwk2->regulareatingwk3) (regulareatingwk3-
>regulareatingwk4) (wklyOBEwk1->wklyOBEwk2 regulareatingwk2) (wklyOBEwk2->wklyOBEwk3
regulareatingwk2 regulareatingwk3) (wklyOBEwk3->wklyOBEwk4 regulareatingwk3
regulareatingwk4) (wklyOBEwk4->regulareatingwk4) (dummy_icontrol mean_bdi dummy_diag-
>regulareatingwk2 regulareatingwk3 regulareatingwk4) (dummy_icontrol everbinge dummy_diag
mean_edewc mean_age->wklyOBEwk2 wklyOBEwk3 wklyOBEwk4),
cov(e.regulareatingwk2*e.regulareatingwk3) cov(e.regulareatingwk3*e.regulareatingwk4)
cov(dummy_icontrol*regulareatingwk1) cov(e.wklyOBEwk2*e.wklyOBEwk3)
cov(e.wklyOBEwk3*e.wklyOBEwk4) cov(dummy_icontrol*wklyOBEwk1) vce(bootstrap, reps(1000)
seed(1242741187)) method(mlmv) standardized
estat bootstrap, bc
```

Appendix 22: Output for study 1c. Alternative Autoregressive structural equation model of the relationships between the intervention (CBT-E), frequency of binge eating as an alternative mediator (BE) and regularity of eating as an alternative outcome (RE)

Model Components	B	Boot. SE	Bias-corrected 95% CI of B		β
			Lower	**Upper**	
BE2→RE2	-0.10	0.04	-0.18	-0.02	-0.23
RE1→RE2	0.58	0.08	0.43	0.75	0.51
BE1→RE2	0.01	0.04	-0.08	0.07	0.01
TX→RE2	1.13	0.17	0.79	1.48	0.38
Depression→RE2	0.00	0.01	-0.02	0.02	0.01
ED diagnosis→RE2	-0.05	0.18	-0.43	0.33	-0.02
RE2→RE3	1.08	0.10	0.89	1.30	0.98
BE2→RE3	0.12	0.04	0.05	0.19	0.23
BE3→RE3	-0.09	0.03	-0.17	-0.03	-0.18
TX→RE3	0.21	0.23	-0.29	0.59	0.06
Depression→RE3	0.00	0.01	-0.02	0.01	-0.02
ED diagnosis→RE3	-0.25	0.16	-0.55	0.06	-0.07
RE3→RE4	0.86	0.07	0.72	0.99	0.89
BE3→RE4	0.05	0.05	-0.03	0.16	0.11
BE4→RE4	-0.10	0.05	-0.21	-0.02	-0.17
TX→RE4	0.36	0.21	-0.07	0.76	0.11
Depression→RE4	0.00	0.01	-0.01	0.02	0.02
ED diagnosis→RE4	0.18	0.15	-0.12	0.46	0.05
BE1→BE2	0.58	0.11	0.36	0.79	0.60
TX→ BE2	-1.03	0.42	-1.88	-0.25	-0.16
ED diagnosis→ BE2	0.92	0.49	-0.02	1.89	0.14
BE Presence→ BE2	0.66	0.33	0.05	1.39	0.08
WC→ BE2	-0.11	0.20	-0.51	0.25	-0.05
Age→BE2	0.00	0.02	-0.05	0.04	-0.01
BE2→BE3	0.74	0.21	0.38	1.16	0.75
TX→BE3	-0.74	0.54	-1.85	0.37	-0.12
ED diagnosis→BE3	0.30	0.68	-1.07	1.58	0.05
BE Presence→BE3	0.02	0.41	-0.78	0.83	0.00
WC →BE3	0.03	0.19	-0.38	0.38	0.01
Age→BE3	-0.01	0.02	-0.05	0.04	-0.02
BE3→BE4	0.80	0.16	0.58	1.22	0.98
TX→BE4	0.66	0.40	-0.09	1.46	0.13
ED diagnosis→BE4	-0.34	0.43	-1.17	0.51	-0.06
BE Presence→BE4	0.03	0.36	-0.78	0.67	0.00
WC →BE4	-0.28	0.19	-0.71	0.04	-0.15
Age→BE4	0.02	0.02	-0.01	0.07	0.07
Var(e.RE2)	0.72	0.09	0.59	0.95	0.32
Var(e.RE3)	0.60	0.12	0.44	0.97	0.23
Var(e.RE4)	0.47	0.09	0.33	0.68	0.19
Var(e.BE2)	4.55	0.74	3.43	6.29	0.44
Var(e.BE3)	4.76	1.11	2.96	7.41	0.47
Var(e.BE4)	4.01	1.38	2.27	9.52	0.59
Cov(e.RE2*e.RE3)	-0.33	0.10	-0.57	-0.16	-0.50

Model Components	B	Boot. SE	Bias-corrected 95% CI of B		β
			Lower	**Upper**	
Cov(e.RE3*e.RE4)	-0.17	0.07	-0.31	-0.05	-0.33
Cov(e.BE2*e.BE3)	-0.76	0.87	-2.62	0.73	-0.16
Cov(e.BE3*e.BE4)	-2.16	1.05	-5.95	-0.85	-0.50
Cov(RE1*BE1)	-1.79	0.43	-2.71	-0.97	-0.41
Cov(RE1*TX)	0.18	0.06	0.07	0.30	0.27
Cov(RE1*Depression)	-3.70	1.31	-6.35	-1.25	-0.26
Cov(RE1*ED diagnosis)	-0.13	0.06	-0.24	-0.01	-0.20
Cov(RE1*BE Presence)	-0.04	0.05	-0.12	0.07	-0.07
Cov(RE1*WC)	0.14	0.18	-0.17	0.54	0.08
Cov(RE1*Age)	0.11	1.03	-2.17	1.92	0.01
Cov(BE1*TX)	-0.13	0.15	-0.44	0.14	-0.08
Cov(BE1*Depression)	0.41	3.59	-6.64	7.14	0.01
Cov(BE1*ED diagnosis)	0.70	0.14	0.44	1.00	0.43
Cov(BE1*BE Presence)	0.43	0.09	0.26	0.64	0.33
Cov(BE1*WC)	-1.09	0.44	-1.97	-0.25	-0.23
Cov(BE1*Age)	-5.08	2.10	-9.21	-0.98	-0.20
Cov(TX*Depression)	-0.37	0.48	-1.41	0.49	-0.07
Cov(TX*ED diagnosis)	-0.01	0.02	-0.05	0.03	-0.05
Cov(TX*BE Presence)	0.01	0.02	-0.02	0.04	0.06
Cov(TX*WC)	0.07	0.06	-0.05	0.20	0.11
Cov(TX*Age)	-0.48	0.31	-1.15	0.13	-0.13
Cov(Depression*ED diagnosis)	0.60	0.47	-0.33	1.46	0.11
Cov(Depression*BE Presence)	0.50	0.35	-0.21	1.18	0.12
Cov(Depression*WC)	4.72	1.49	1.96	7.96	0.31
Cov(Depression*Age)	13.13	7.95	-2.76	28.72	0.16
Cov(ED diagnosis*BE Presence)	0.08	0.01	0.06	0.11	0.40
Cov(ED diagnosis*WC)	-0.02	0.06	-0.15	0.08	-0.03
Cov(ED diagnosis*Age)	-0.37	0.32	-0.99	0.25	-0.10
Cov(BE Presence*WC)	0.01	0.05	-0.08	0.10	0.01
Cov(BE Presence*Age)	-0.27	0.30	-0.94	0.26	-0.09
Cov(WC*Age)	0.45	1.10	-2.43	2.16	0.04
Successful bootstraps	1000/1000				

Notes. B=unstandardized coefficient. SE=standard error. CI=confidence interval. β=standardised coefficient. RE# = regularity of eating where # indicates intervention week. BE# = frequency of binge eating where # indicates intervention week. TX=treatment allocation, BE Presence=Presence of binge eating in the month prior to starting treatment, WC = weight concern at baseline, ED diagnosis = eating disorder diagnosis (bulimia nervosa vs. eating disorder not otherwise specified)

Appendix 23: Stata code for study 1d latent growth curve model

```
****CODE FOR LATENT GROWTH CURVE MODEL*****
- trialno = patient ID
- dummy_icontrol = treatment allocation where 0 is IPT-ED, and 1 is CBT-E
- Ire = intercept of regularity of eating (score at baseline)
- Ibe = intercept of frequency of binge eating (score at baseline)
- Sre = slope of regularity of eating (change score)
- Sbe = slope of frequency of binge eating (change score)
- regulareatingwk1 = regularity of eating at week 1
- regulareatingwk2 = regularity of eating at week 2
- regulareatingwk3 = regularity of eating at week 3
- regulareatingwk4 = regularity of eating at week 4
- wklyOBEwk1 = frequency of binge eating at week 1
- wklyOBEwk2 = frequency of binge eating at week 2
- wklyOBEwk3 = frequency of binge eating at week 3
- wklyOBEwk4 = frequency of binge eating at week 4
- dummy_diag = eating disorder diagnosis (bulimia nervosa vs. eating disorder not otherwise
specified)
- mean_edewc = level of  weight concern at baseline centred at the sample mean
- everbinge = presence of binge eating in the month prior to treatment
- mean_age = age at baseline centred at the sample mean
- mean_bdi = depression at baseline centred at the sample mean
*/


*unstandardised model
sem (_cons-> dummy_icontrol Ire Sre Ibe Sbe mean_bdi dummy_diag mean_edewc dummy_diag
everbinge mean_age) (dummy_icontrol-> Sre Sbe) (Ire->regulareatingwk1@1 regulareatingwk2@1
regulareatingwk3@1 regulareatingwk4@1 Sbe) (Sre->regulareatingwk1@0 regulareatingwk2@.80
regulareatingwk3@1 regulareatingwk4@1 Sbe) (Ibe->wklyOBEwk1@1 wklyOBEwk2@1
wklyOBEwk3@1 wklyOBEwk4@1) (Sbe->wklyOBEwk1@0 wklyOBEwk2@.2 wklyOBEwk3@.5
wklyOBEwk4@1) (mean_bdi dummy_diag-> Sre) (mean_edewc dummy_diag everbinge mean_age->
Sbe), cov(e.Ire*e.Ibe) cov(e.Ire*e.Sre) cov(e.Ibe*e.Sbe) covstruct(e.Ire e.Ibe e.dummy_icontrol
e.mean_bdi e.mean_edewc e.dummy_diag e.everbinge e.mean_age, uns) nocons method(mlmv)
allmissing vce(bootstrap, reps(1000) seed(644045765))
estat boot, bc
estat gof, stats(all)

*run bootstrap on manual calculations of indirect and total effects for treatment
program indireff, rclass
sem (_cons-> dummy_icontrol Ire Sre Ibe Sbe mean_bdi dummy_diag mean_edewc dummy_diag
everbinge mean_age) (dummy_icontrol-> Sre Sbe) (Ire->regulareatingwk1@1 regulareatingwk2@1
regulareatingwk3@1 regulareatingwk4@1 Sbe) (Sre->regulareatingwk1@0 regulareatingwk2@.80
regulareatingwk3@1 regulareatingwk4@1 Sbe) (Ibe->wklyOBEwk1@1 wklyOBEwk2@1
wklyOBEwk3@1 wklyOBEwk4@1) (Sbe->wklyOBEwk1@0 wklyOBEwk2@.2 wklyOBEwk3@.5
wklyOBEwk4@1) (mean_bdi dummy_diag-> Sre) (mean_edewc dummy_diag everbinge mean_age->
Sbe), cov(e.Ire*e.Ibe) cov(e.Ire*e.Sre) cov(e.Ibe*e.Sbe) covstruct(e.Ire e.Ibe e.dummy_icontrol
e.mean_bdi e.mean_edewc e.dummy_diag e.everbinge e.mean_age, uns) nocons method(mlmv)
allmissing
mat b = r(table)
display b[1,8]*b[1,19]
display b[1,8]*b[1,19]+b[1,13]
return scalar bi  = b[1,8]*b[1,19]
return scalar bt  = b[1,8]*b[1,19]+b[1,13]
```

```
end

bootstrap r(bi) r(bt), reps(1000) seed(644045765): indireff
estat bootstrap, bc
program drop indireff

*standardised model
sem (_cons-> dummy_icontrol Ire Sre Ibe Sbe mean_bdi dummy_diag mean_edewc dummy_diag
everbinge mean_age) (dummy_icontrol-> Sre Sbe) (Ire->regulareatingwk1@1 regulareatingwk2@1
regulareatingwk3@1 regulareatingwk4@1 Sbe) (Sre->regulareatingwk1@0 regulareatingwk2@.80
regulareatingwk3@1 regulareatingwk4@1 Sbe) (Ibe->wklyOBEwk1@1 wklyOBEwk2@1
wklyOBEwk3@1 wklyOBEwk4@1) (Sbe->wklyOBEwk1@0 wklyOBEwk2@.2 wklyOBEwk3@.5
wklyOBEwk4@1) (mean_bdi dummy_diag-> Sre) (mean_edewc dummy_diag everbinge mean_age->
Sbe), cov(e.Ire*e.Ibe) cov(e.Ire*e.Sre) cov(e.Ibe*e.Sbe) covstruct(e.Ire e.Ibe e.dummy_icontrol
e.mean_bdi e.mean_edewc e.dummy_diag e.everbinge e.mean_age, uns) nocons method(mlmv)
allmissing vce(bootstrap, reps(1000) seed(644045765)) standardized
```

Appendix 24: Output for study 1d. Parallel process latent growth curve model of the relationships between the intervention (CBT-E), change in the purported mediator (regularity of eating; RE) and change in the outcome (frequency of binge eating; BE)

Model Components	B	Boot. SE	Bias-corrected 95% CI of B		β
			Lower	**Upper**	
Baseline RE (intercept)	2.85	0.12	2.60	3.08	2.62
TX→Change in RE (slope)	1.36	0.19	0.97	1.74	0.85
Depression→Change in RE (slope)	0.01	0.01	-0.01	0.03	0.16
ED diagnosis→Change in RE (slope)	-0.20	0.20	-0.51	0.22	-0.12
Constant→Change in RE (slope)	0.17	0.11	-0.09	0.37	0.22
Baseline BE (intercept)	2.63	0.30	2.12	3.28	0.90
TX→ Change in BE (slope)	2.42	2.04	-0.11	6.94	0.53
ED diagnosis→ Change in BE (slope)	-1.20	0.75	-2.36	0.51	-0.26
WC→ Change in BE (slope)	-0.04	0.21	-0.57	0.29	-0.03
BE Presence→ Change in BE (slope)	-0.58	0.44	-1.59	0.14	-0.10
Age→ Change in BE (slope)	0.06	0.02	0.01	0.10	0.18
Baseline RE (intercept)→ Change in BE (slope)	0.44	0.46	-0.20	1.40	0.21
Change in RE (slope) → Change in BE (slope)	-1.80	1.72	-6.48	-0.02	-0.63
Constant→ Change in BE (slope)	-1.15	1.40	-4.25	0.80	-0.50
Baseline RE (intercept) →RE1	1.00	-	-	-	0.83
Baseline RE (intercept) →RE2	1.00	-	-	-	0.72
Change in RE (slope) →RE2	0.80	-	-	-	0.42
Baseline RE (intercept) →RE3	1.00	-	-	-	0.68
Change in RE (slope) →RE3	1.00	-	-	-	0.50
Baseline RE (intercept) →RE4	1.00	-	-	-	0.68
Change in RE (slope) →RE4	1.00	-	-	-	0.50
Baseline BE (intercept) →BE1	1.00	-	-	-	0.85
Baseline BE (intercept) →BE2	1.00	-	-	-	0.94
Change in BE (slope) →BE2	0.20	-	-	-	0.15
Baseline BE (intercept) →BE3	1.00	-	-	-	0.92
Change in BE (slope) →BE3	0.50	-	-	-	0.36
Baseline BE (intercept) →BE4	1.00	-	-	-	1.12
Change in BE (slope) →BE4	1.00	-	-	-	0.87
Var(e.RE1)	0.55	0.17	0.24	0.88	0.32
Var(e.RE2)	0.36	0.06	0.24	0.49	0.16
Var(e.RE3)	0.30	0.07	0.19	0.48	0.12
Var(e.RE4)	0.30	0.06	0.19	0.46	0.12
Var(e.BE1)	3.44	0.99	1.39	4.96	0.29
Var(e.BE2)	2.30	0.51	1.33	3.35	0.24
Var(e.BE3)	3.81	1.11	2.20	6.27	0.38
Var(e.BE4)	0.33	0.60	0.00	1.11	0.05
Var(e.baseline RE (intercept))	1.19	0.25	0.62	1.61	1.00
Var(e.change in RE (slope))	0.16	0.16	0.02	0.61	0.25
Var(e.baseline BE (intercept))	8.61	1.65	5.76	12.55	1.00
Var(e.change in BE (slope))	3.74	1.43	2.00	6.61	0.71
Cov(e.TX*e.Depression)	-0.39	0.44	-1.30	0.37	-0.07
Cov(e.TX*e.ED diagnosis)	-0.01	0.02	-0.05	0.03	-0.05

Model Components	B	Boot. SE	Bias-corrected 95% CI of B		β
			Lower	**Upper**	
Cov(e.TX*e.WC)	0.07	0.06	-0.07	0.19	0.11
Cov(e.TX*e.BE Presence)	0.01	0.02	-0.02	0.05	0.06
Cov(e.TX*e.Age)	-0.48	0.34	-1.19	0.12	-0.13
Cov(e.TX*e.baseline RE (intercept))	0.17	0.06	0.06	0.29	0.32
Cov(e.TX*e.baseline BE (intercept))	-0.28	0.14	-0.52	0.03	-0.19
Cov(e.Depression*e.ED diagnosis)	0.61	0.46	-0.46	1.48	0.11
Cov(e.Depression*e.WC)	4.71	1.40	2.35	7.84	0.31
Cov(e.Depression*e.BE Presence)	0.50	0.33	-0.24	1.09	0.12
Cov(e.Depression*e.Age)	13.30	7.79	-1.35	28.57	0.16
Cov(e.Depression*e.baseline RE (intercept))	-3.60	1.35	-6.26	-0.93	-0.31
Cov(e.Depression*e.baseline BE (intercept))	2.46	3.07	-4.99	8.22	0.08
Cov(e.ED diagnosis*e.WC)	-0.02	0.06	-0.15	0.09	-0.03
Cov(e.ED diagnosis*e.BE Presence)	0.08	0.01	0.05	0.10	0.40
Cov(e.ED diagnosis*e.Age)	-0.37	0.30	-0.97	0.25	-0.10
Cov(e.ED diagnosis*e.baseline RE (intercept))	-0.13	0.06	-0.24	-0.02	-0.24
Cov(e.ED diagnosis*e.baseline BE (intercept))	0.74	0.14	0.47	1.06	0.51
Cov(e.WC*e.BE Presence)	0.01	0.04	-0.07	0.10	0.01
Cov(e.WC*e.Age)	0.45	1.03	-1.74	2.29	0.04
Cov(e.WC *e.baseline RE (intercept))	0.14	0.17	-0.19	0.46	0.09
Cov(e.WC *e.baseline BE (intercept))	-0.96	0.40	-1.68	-0.12	-0.23
Cov(e.BE Presence*e.Age)	-0.27	0.28	-0.93	0.17	-0.09
Cov(e.BE Presence*e.baseline RE (intercept))	-0.03	0.04	-0.12	0.06	-0.06
Cov(e.BE Presence*e.baseline BE (intercept))	0.44	0.09	0.28	0.63	0.38
Cov(e.Age*e.baseline RE (intercept))	0.31	0.95	-1.87	1.90	0.04
Cov(e.Age*e.baseline BE (intercept))	-4.46	2.06	-8.17	-0.52	-0.20
Cov(e.baseline RE (intercept)*e.change in RE (slope))	0.00	0.15	-0.26	0.19	0.01
Cov(e.baseline RE (intercept)*e.baseline BE (intercept)*)	-1.74	0.43	-2.69	-1.05	-0.54
Cov(e.baseline BE (intercept)*e.change in BE (slope))	-1.75	0.73	-3.55	-0.68	-0.31
Successful bootstraps	350/1000				

Notes. B=unstandardized coefficient. SE=standard error. CI=confidence interval. β=standardised coefficient. RE# = regularity of eating where # indicates intervention week. BE# = frequency of binge eating where # indicates intervention week. TX=treatment allocation, BE Presence=Presence of binge eating in the month prior to starting treatment, WC = weight concern at baseline.

Statistical Output and Stata Code for Study 2: Frequency of Weighing as a Mediator of the Effect of Enhanced Cognitive Behavioural Therapy on Concern about Weight

Appendix 25: Stata code for study 2a multilevel models and their coefficient plots

```
**********CODE FOR FINAL MULTILEVEL MODELS**********
/*
Notes on variables.
- trialno = patient ID
- dummy_icontrol = treatment allocation where 0 is IPT-ED, and 1 is CBT-E
- adj_eqq16 = frequency of weighing
- center_ eqq16 = frequency of weighing (centred around sample mean at baseline)
- eqq26 = concern about weight
- dummy_txweek = treatment week, where zero indicates baseline (start of treatment)
- dummy_txweek2 = treatment week as a quadratic trend
- frequentSW = frequent weigher status where 0 is non-frequent weigher, and 1 is frequent weigher
- rsepre30cat = self-esteem group where 0 is low self-esteem, and 1 is high self-esteem
- freqweighXtx = interaction between frequent weigher status and treatment allocation
- rsepre30Xtx = interaction between self-esteem group and treatment allocation
- mean_edewc = level of  weight concern at baseline centred at the sample mean
- mean_rsepre30 = level of self-esteem at baseline centred at the sample mean
*/

*****STEP 1- IMPACT OF THE INTERVENTION ON THE PURPORTED MEDIATOR*****
*run final model
bootstrap, cluster(trialno) reps(1000) seed(2076418861): xtmixed adj_eqq16 c.dummy_txweek
c.dummy_txweek2 dummy_icontrol frequentSW rsepre30cat
i.dummy_icontrol#i.frequentSW#c.dummy_txweek
i.dummy_icontrol#i.frequentSW#c.dummy_txweek2 rsepre30cat#dummy_icontrol
c.dummy_txweek#c.mean_edewc c.dummy_txweek2#c.mean_edewc || trialno: dummy_txweek if
txweek<7, reml cov(unstr) var
estat boot, bc
*create summary table of model
eststo est1
estout est1, cells("b(fmt(2)) se p(fmt(3)) ci_bc[1] ci_bc[2]") modelwidth(9) style(fixed) stats(aic
N_clust N N_reps)
*create coefficient plot of key variables
coefplot est1, omitted baselevels keep(dummy_txweek dummy_txweek2
1.dummy_icontrol#0.frequentSW#?.dummy_txweek
1.dummy_icontrol#1.frequentSW#?.dummy_txweek
1.dummy_icontrol#1.frequentSW#?.dummy_txweek2
0.dummy_icontrol#1.frequentSW#?.dummy_txweek) xline(0) ci(ci_bc) ciopts(recast(rcap)) mlabel
format(%9.2f) mlabposition(12) mlabgap(*1) name(FWMLMcoeffinterval, replace) xtitle("{bf:Effect
on Frequency of Weighing}", height(5) ) title("{bf:1. Impact of Intervention on Purported Mediator}",
size(medlarge) color(black) span) coeflabel(dummy_txweek = "Treatment week"  dummy_txweek2 =
"Treatment week (quadratic trend)" 1.dummy_icontrol#1.frequentSW#?.dummy_txweek ="Freq.
weigher - CBT-E (linear trend)" 1.dummy_icontrol#1.frequentSW#?.dummy_txweek2 ="Freq. weigher
- CBT-E (quadratic trend)" 0.dummy_icontrol#1.frequentSW#?.dummy_txweek = "Freq. weigher - IPT-
```

```
ED (linear trend)" 1.dummy_icontrol#0.frequentSW#?.dummy_txweek = "Non-freq. weigher - CBT-E
(linear trend)", wrap(43)) order (dummy_txweek dummy_txweek2
1.dummy_icontrol#0.frequentSW#?.dummy_txweek
0.dummy_icontrol#1.frequentSW#?.dummy_txweek
0.dummy_icontrol#1.frequentSW#?.dummy_txweek2) headings(dummy_txweek  = "{bf:Impact of
Treatment Week}"1.dummy_icontrol#0.frequentSW#?.dummy_txweek  = "{bf:Impact of Treatment
Allocation}") saving(FWMLMcoeffinterval, replace)

****STEP 2 - IMPACT OF THE INTERVENTION ON THE OUTCOME****
*run final model
bootstrap, cluster(trialno) reps(1000) seed(2076418861): xtmixed eqq26 c.dummy_txweek
c.dummy_txweek2 dummy_icontrol c.dummy_txweek#dummy_icontrol mean_edewc
c.mean_edewc#c.mean_edewc mean_rsepre30 c.mean_rsepre30#c.mean_rsepre30 || trialno:
dummy_txweek if txweek<7, reml cov(unstr) var
estat boot, bc
estat icc
estat ic
*create summary table of model
eststo est2
estout est2, cells("b(fmt(2)) se p(fmt(3)) ci_bc[1] ci_bc[2]") modelwidth(9) style(fixed) stats(aic
N_clust N N_reps)
*create coefficient plot of key variables
coefplot est2, omitted baselevels keep(dummy_txweek dummy_txweek2
1.dummy_icontrol#c.dummy_txweek ) xline(0) ci(ci_bc) ciopts(recast(rcap)) mlabel format(%9.2f)
mlabposition(12) mlabgap(*1) name(CWMLMcoeffinterval, replace) xtitle("{bf:Effect on Concern
about Weight}", height(5)) title("{bf:2. Impact of Intervention on Outcome}", size(medlarge)
color(black) span) coeflabel(dummy_txweek = "Treatment week" dummy_txweek2 = "Treatment
week (quadratic trend)" 1.dummy_icontrol#?.dummy_txweek = "CBT-E (linear trend)", wrap(40))
order (dummy_txweek dummy_txweek2 1.dummy_icontrol#dummy_txweek)
headings(dummy_txweek  = "{bf:Impact of Treatment Week}" 1.dummy_icontrol#?.dummy_txweek
= "{bf:Impact of Treatment Allocation}") saving(CWMLMcoeffinterval, replace)

****STEP 3 - IMPACT OF THE PURPORTED MEDIATOR ON THE OUTCOME****
*run final model
bootstrap, cluster(trialno) reps(1000) seed(2076418861): xtmixed eqq26 c.dummy_txweek
c.dummy_txweek2 center_adj_eqq16 mean_edewc c.mean_edewc#c.mean_edewc mean_rsepre30
c.mean_rsepre30#c.mean_rsepre30 || trialno: dummy_txweek if txweek<7, reml cov(unstr) var
estat boot, bc
estat icc
estat ic
*create summary table of model
eststo est3
estout est3, cells("b(fmt(2)) se p(fmt(3)) ci_bc[1] ci_bc[2]") modelwidth(9) style(fixed) stats(aic
N_clust N N_reps)
*create coefficient plot of key variables
coefplot est3, omitted baselevels keep(dummy_txweek dummy_txweek2 center_adj_eqq16) xline(0)
ci(ci_bc) ciopts(recast(rcap)) mlabel format(%9.2f) mlabposition(12) mlabgap(*1)
name(MOFWMLMcoeffinterval, replace) xtitle("{bf:Effect on Concern about Weight}", height(5))
title("{bf:3. Impact of Purported Mediator on Outcome}", size(medlarge) color(black) span)
coeflabel(dummy_txweek = "Treatment week" dummy_txweek2 = "Treatment week (quadratic
trend)" center_adj_eqq16 = "Frequency of weighing", wrap(40)) order (dummy_txweek
dummy_txweek2 center_adj_eqq16) headings(dummy_txweek  = "{bf:Impact of Treatment Week}"
center_adj_eqq16  = "{bf:Impact of Purported Mediator}") saving(MOFWMLMcoeffinterval, replace)
```

```
****STEP 4 - IMPACT OF THE INTERVENTION ON THE OUTCOME WHEN INCLUDING THE PURPORTED MEDIATOR IN THE MODEL (CONTEMPORANEOUS RELATIONSHIP)****
*run final model
bootstrap, cluster(trialno) reps(5) seed(2076418861): xtmixed eqq26 c.dummy_txweek c.dummy_txweek2 dummy_icontrol c.dummy_txweek#dummy_icontrol center_adj_eqq16 mean_edewc c.mean_edewc#c.mean_edewc mean_rsepre30 c.mean_rsepre30#c.mean_rsepre30 || trialno: dummy_txweek if txweek<7, reml cov(unstr) var
estat boot, bc
estat icc
estat ic
*create summary table of model
eststo est4
estout est4, cells("b(fmt(2)) se p(fmt(3)) ci_bc[1] ci_bc[2]") modelwidth(9) style(fixed) stats(aic N_clust N N_reps)
*create coefficient plot of key variables
coefplot est4, omitted baselevels keep(dummy_txweek dummy_txweek2 1.dummy_icontrol#c.dummy_txweek center_adj_eqq16) xline(0) ci(ci_bc) ciopts(recast(rcap)) mlabel format(%9.2f) mlabposition(12) mlabgap(*1) name(medFWMLMcoeffinterval, replace) xtitle("{bf:Effect on Concern about Weight}", height(5) ) title("{bf:4. Impact of Intervention on Outcome with Purported Mediator}", size(medlarge) color(black) span) coeflabel(dummy_txweek = "Treatment week" dummy_txweek2 = "Treatment week (quadratic trend)" 1.dummy_icontrol#?.dummy_txweek = "CBT-E (linear trend)" center_adj_eqq16 = "Frequency of weighing", wrap(40)) order (dummy_txweek dummy_txweek2 1.dummy_ccontrol#?.dummy_txweek center_adj_eqq16)  headings(dummy_txweek  = "{bf:Impact of Treatment Week}" 1.dummy_icontrol#?.dummy_txweek  = "{bf:Impact of Treatment Allocation}" center_adj_eqq16 = "{bf:Impact of Purported Mediator}") saving(medFWMLMcoeffinterval, replace)

****COMBINE COEFPLOTS INTO ONE GRAPH****
graph combine FWMLMcoeffinterval CWMLMcoeffinterval MOFWMLMcoeffinterval medFWMLMcoeffinterval, name(combFW_MLM, replace) xsize(20) ysize(10) saving(combFW_MLM, replace)
```

Appendix 26: Output for study 2a, step 1. Multi-level model of the impact of the intervention (CBT-E) on the purported mediator (frequency of weighing)

Model Components	Variable Details	B	Boot. SE	p	Bias-corrected 95% CI of B	
					Lower	**Upper**
Baseline predictors	Treatment allocation (CBT-E vs. IPT-ED)	0.68	0.70	.328	-0.67	1.91
	Frequent weigher status (frequent vs. non-frequent weigher)	6.19	0.85	.000	4.49	7.67
	Self-esteem (high vs. low self-esteem)	-1.80	0.76	.017	-3.25	-0.31
Impact of treatment week	Treatment week (linear trend)	0.04	0.24	.854	-0.46	0.50
	Treatment week (quadratic)	0.02	0.06	.694	-0.09	0.14
Impact of baseline predictors during treatment	Weight concern (linear trend)	-0.28	0.12	.017	-0.53	-0.08
	Weight concern (quadratic trend)	0.06	0.03	.012	0.02	0.12
Moderators of treatment	Low self-esteem in CBT-E	-1.70	0.76	.027	-3.07	-0.23
	Non-frequent weigher in CBT-E*treatment week (linear trend)	0.48	0.28	.094	-0.05	1.06
	Non-frequent weigher in CBT-E*treatment week (quadratic trend)	-0.10	0.06	.094	-0.23	0.02
	Frequent weigher in IPT-ED*treatment week (linear trend)	-0.51	0.48	.287	-1.40	0.46
	Frequent weigher in IPT-ED*treatment week (quadratic trend)	-0.05	0.10	.574	-0.24	0.13
	Frequent weigher in CBT-E*treatment week (linear trend)	-3.25	0.91	.000	-5.00	-1.56
	Frequent weigher in CBT-E*treatment week (quadratic trend)	0.41	0.14	.004	0.16	0.72
Constant	Non-frequent weighers in IPT-ED with low self-esteem and average weight concern at baseline	1.49	0.64	.020	0.28	2.75
Random effects (log values)	SD (treatment week)	-0.16	0.29	.582	-1.03	0.08
	SD (constant)	1.42	0.21	.000	0.74	1.63
	Correlation (treatment week*constant)	-1.05	0.35	.003	-1.83	-0.42
	SD (residual)	0.54	0.14	.000	0.39	0.75
Statistics	AIC	3115.80				
	Number of patients	114				
	Number of observations	658				
	Number of bootstraps	1000/1000				

Notes. B=unstandardized coefficient, p-value based on 95% confidence interval, CI=confidence interval.

Appendix 27: Output for study 2a, step 2. Multi-level model of the impact of the intervention (CBT-E) on the outcome (concern about weight) It would be good to see the estimates of the SD of the random effects and of the residual errors

Model Components	Variable Details	B	Boot. SE	p	Bias-corrected 95% CI of B	
					Lower	**Upper**
Baseline predictors	Treatment allocation (CBT-E vs. IPT-ED)	0.00	0.19	.980	-0.39	0.37
	Weight concern	0.68	0.06	.000	0.55	0.80
	Weight concern (quadratic)	-0.10	0.03	.001	-0.16	-0.03
	Self-esteem	-0.02	0.03	.385	-0.07	0.03
	Self-esteem (quadratic)	0.02	0.01	.011	0.00	0.03
Impact of treatment week	Treatment week (linear trend)	-0.36	0.07	.000	-0.52	-0.22
	Treatment week (quadratic)	0.03	0.01	.015	0.01	0.06
Impact of treatment allocation	CBT-E*treatment week (linear trend)	0.04	0.05	.434	-0.06	0.13
Constant	Patients in IPT-ED with average weight concern and self-esteem at baseline	4.61	0.18	.000	4.27	4.96
Random effects (log values)	SD (treatment week)	-1.27	0.09	.000	-1.53	-1.19
	SD (constant)	0.12	0.08	.146	-0.16	0.23
	Correlation (treatment week*constant)	-0.26	0.10	.011	-0.37	0.09
	SD (residual)	-0.13	0.05	.009	-0.13	-0.08
Statistics	AIC	2321.72				
	Number of patients	125				
	Number of observations	718				
	Number of bootstraps	1000/1000				

Notes. B=unstandardized coefficient, *p*-value based on 95% confidence interval, CI=confidence interval.

Appendix 28: Output for study 2a, step 3. Multi-level model of impact of the purported mediator (frequency of weighing) on the outcome (concern about weight)

Model Components	Variable Details	B	Boot. SE	p	Bias-corrected 95% CI of B	
					Lower	**Upper**
Baseline predictors	Weight concern	0.71	0.07	.000	0.58	0.84
	Weight concern (quadratic)	-0.09	0.03	.005	-0.14	-0.02
	Self-esteem	-0.02	0.02	.420	-0.07	0.03
	Self-esteem (quadratic)	0.01	0.01	.012	0.00	0.02
Impact of treatment week	Treatment week (linear trend)	-0.33	0.07	.000	-0.49	-0.20
	Treatment week (quadratic)	0.03	0.01	.017	0.01	0.06
Impact of purported mediator	Frequency of weighing	0.04	0.02	.047	0.00	0.07
Constant	Patients with average weight concern and self-esteem at baseline	4.61	0.14	.000	4.33	4.88
Random effects (log values)	SD (treatment week)	-1.31	0.09	.000	-1.55	-1.22
	SD (constant)	0.12	0.08	.148	-0.15	0.23
	Correlation (treatment week*constant)	-0.24	0.10	.015	-0.35	0.09
	SD (residual)	-0.13	0.05	.009	-0.12	-0.08
Model statistics	AIC	2306.72				
	Number of patients	125				
	Number of observations	714				
	Number of bootstraps	1000/1000				

Notes. B=unstandardized coefficient, *p*-value based on 95% confidence interval, CI=confidence interval.

Appendix 29: Output for study 2a, step 4. Multi-level model of impact of the intervention (CBT-E) on the outcome (concern about weight) when accounting for the purported mediator (frequency of weighing)

Model Components	Variable Details	B	Boot. SE	p	Bias-corrected 95% CI of B	
					Lower	**Upper**
Baseline predictors	Treatment allocation (CBT-E vs. IPT-ED)	0.02	0.19	.935	-0.37	0.39
	Weight concern	0.71	0.07	.000	0.58	0.84
	Weight concern (quadratic)	-0.09	0.03	.006	-0.14	-0.02
	Self-esteem	-0.02	0.03	.383	-0.07	0.03
	Self-esteem (quadratic)	0.01	0.01	.013	0.00	0.02
Impact of treatment week	Treatment week (linear trend)	-0.36	0.08	.000	-0.51	-0.21
	Treatment week (quadratic)	0.03	0.01	.017	0.01	0.06
Impact of treatment allocation	CBT-E*treatment week (linear trend)	0.05	0.05	.337	-0.06	0.14
Impact of purported mediator	Frequency of weighing	0.04	0.02	.044	0.00	0.07
Constant	Patients in IPT-ED with average weight concern and self-esteem at baseline	4.59	0.18	.000	4.25	4.94
Random effects (log values)	SD (treatment week)	-1.30	0.09	.000	-1.54	-1.21
	SD (constant)	0.13	0.08	.132	-0.15	0.24
	Correlation (treatment week*constant)	-0.25	0.10	.016	-0.36	0.11
	SD (residual)	-0.13	0.05	.009	-0.12	-0.08
Model statistics	AIC	2314.91				
	Number of patients	125				
	Number of observations	714				
	Number of bootstraps	1000/1000				

Notes. B=unstandardized coefficient, *p*-value based on 95% confidence interval, CI=confidence interval.

Appendix 30: Stata code for study 2b autoregressive structural equation model

```
*******CODE FOR FINAL AUTOREGRESSIVE STRUCTURAL EQUATION MODEL*******
/*
Notes on variables.
- trialno = patient ID
- dummy_icontrol = treatment allocation where 0 is IPT-ED, and 1 is CBT-E
- adj_eqq16wk1 = frequency of weighing at week 1
- adj_eqq16wk2 = frequency of weighing at week 2
- adj_eqq16wk3 = frequency of weighing at week 3
- adj_eqq16wk4 = frequency of weighing at week 4
- adj_eqq16wk5 = frequency of weighing at week 5
- adj_eqq16wk6 = frequency of weighing at week 6
- eqq26wk1 = concern about weight at week 1
- eqq26wk2 = concern about weight at week 2
- eqq26wk3 = concern about weight at week 3
- eqq26wk4 = concern about weight at week 4
- eqq26wk5 = concern about weight at week 5
- eqq26wk6 = concern about weight at week 6
- frequentSW = frequent weigher status where 0 is non-frequent weigher, and 1 is frequent weigher
- rsepre30cat = self-esteem group where 0 is low self-esteem, and 1 is high self-esteem
- freqweighXtx = interaction between frequent weigher status and treatment allocation
- rsepre30Xtx = interaction between self-esteem group and treatment allocation
- mean_edewc = level of  weight concern at baseline (centred at the sample mean)
- mean_edewc2 = level of  weight concern at baseline (centred at the sample mean) – fitted as a
quadratic
- mean_rsepre30 = level of self-esteem at baseline (centred at the sample mean)
- mean_rsepre302 = level of self-esteem at baseline (centred at the sample mean) – fitted as a
quadratic
*/

*unstandardised
sem (adj_eqq16wk1->adj_eqq16wk2 eqq26wk2) (adj_eqq16wk2->adj_eqq16wk3 eqq26wk2
eqq26wk3) (adj_eqq16wk3->adj_eqq16wk4 eqq26wk3 eqq26wk4) (adj_eqq16wk4->adj_eqq16wk5
eqq26wk4 eqq26wk5) (adj_eqq16wk5->adj_eqq16wk6 eqq26wk5 eqq26wk6) (adj_eqq16wk6->
eqq26wk6)(eqq26wk1->eqq26wk2) (eqq26wk2->eqq26wk3) (eqq26wk3->eqq26wk4) (eqq26wk4-
>eqq26wk5) (eqq26wk5->eqq26wk6) (dummy_icontrol frequentSW rsepre30cat freqweighXtx
rsepre30Xtx->adj_eqq16wk2 adj_eqq16wk3 adj_eqq16wk4 adj_eqq16wk5 adj_eqq16wk6)
(dummy_icontrol  mean_edewc mean_rsepre30 mean_edewc2 mean_rsepre302 -> eqq26wk2
eqq26wk3 eqq26wk4 eqq26wk5 eqq26wk6), cov(e.adj_eqq16wk2*e.adj_eqq16wk3)
cov(e.adj_eqq16wk3*e.adj_eqq16wk4) cov(e.adj_eqq16wk4*e.adj_eqq16wk5)
cov(e.adj_eqq16wk5*e.adj_eqq16wk6) cov(e.eqq26wk2*e.eqq26wk3) cov(e.eqq26wk3*e.eqq26wk4)
cov(e.eqq26wk4*e.eqq26wk5) cov(e.eqq26wk5*e.eqq26wk6) cov(_Ex) vce(bootstrap, reps(1000)
seed(1552470950)) method(mlmv) allmissing
estat boot, bc
*fit indices
estat gof, stats(all)
*test stability of path coefficients for frequency of weighing
test(_b[adj_eqq16wk2:adj_eqq16wk1]==_b[adj_eqq16wk3:adj_eqq16wk2]==_b[adj_eqq16wk4:adj_
eqq16wk3]==_b[adj_eqq16wk5:adj_eqq16wk4]==_b[adj_eqq16wk6:adj_eqq16wk5])
*test stability of path coefficients for concern about weight
test(_b[eqq26wk2:eqq26wk1]==_b[eqq26wk3:eqq26wk2]==_b[eqq26wk4:eqq26wk3]==_b[eqq26wk
5:eqq26wk4]==_b[eqq26wk6:eqq26wk5])
*get table of direct, indirect and total effects of treatment, set matrices & return scalars for model to
run bootstrap teffects for effect of CBT-E
```

```
estat teffects, standardized
mat list r(indirect)
mat list r(total)
*In non-frequent weighers
program indireff_NFW, rclass
set more off
sem (adj_eqq16wk1->adj_eqq16wk2 eqq26wk2) (adj_eqq16wk2->adj_eqq16wk3 eqq26wk2
eqq26wk3) (adj_eqq16wk3->adj_eqq16wk4 eqq26wk3 eqq26wk4) (adj_eqq16wk4->adj_eqq16wk5
eqq26wk4 eqq26wk5) (adj_eqq16wk5->adj_eqq16wk6 eqq26wk5 eqq26wk6) (adj_eqq16wk6->
eqq26wk6) (eqq26wk1->eqq26wk2) (eqq26wk2->eqq26wk3) (eqq26wk3->eqq26wk4) (eqq26wk4-
>eqq26wk5) (eqq26wk5->eqq26wk6) (dummy_icontrol frequentSW rsepre30cat freqweighXtx
rsepre30Xtx->adj_eqq16wk2 adj_eqq16wk3 adj_eqq16wk4 adj_eqq16wk5 adj_eqq16wk6)
(dummy_icontrol  mean_edewc mean_rsepre30 mean_edewc2 mean_rsepre302 -> eqq26wk2
eqq26wk3 eqq26wk4 eqq26wk5 eqq26wk6), cov(e.adj_eqq16wk2*e.adj_eqq16wk3)
cov(e.adj_eqq16wk3*e.adj_eqq16wk4) cov(e.adj_eqq16wk4*e.adj_eqq16wk5)
cov(e.adj_eqq16wk5*e.adj_eqq16wk6) cov(e.eqq26wk2*e.eqq26wk3) cov(e.eqq26wk3*e.eqq26wk4)
cov(e.eqq26wk4*e.eqq26wk5) cov(e.eqq26wk5*e.eqq26wk6) cov(_Ex) method(mlmv) allmissing
estat teffects
mat bi = r(indirect)
mat bt = r(total)
return scalar biwk2  = bi[1,10]
return scalar biwk3  = bi[1,31]
return scalar biwk4  = bi[1,55]
return scalar biwk5  = bi[1,82]
return scalar biwk6  = bi[1,112]
return scalar btwk2  = bt[1,10]
return scalar btwk3  = bt[1,31]
return scalar btwk4  = bt[1,55]
return scalar btwk5  = bt[1,82]
return scalar btwk6  = bt[1,112]
end

bootstrap r(biwk2) r(biwk3) r(biwk4) r(biwk5) r(biwk6) r(btwk2) r(btwk3) r(btwk4) r(btwk5) r(btwk6),
reps(1000) seed(1552470950): indireff_NFW
estat bootstrap, bc
program drop indireff_NFW

*In frequent weighers
program indireff_FW, rclass
set more off
sem (adj_eqq16wk1->adj_eqq16wk2 eqq26wk2) (adj_eqq16wk2->adj_eqq16wk3 eqq26wk2
eqq26wk3) (adj_eqq16wk3->adj_eqq16wk4 eqq26wk3 eqq26wk4) (adj_eqq16wk4->adj_eqq16wk5
eqq26wk4 eqq26wk5) (adj_eqq16wk5->adj_eqq16wk6 eqq26wk5 eqq26wk6) (adj_eqq16wk6->
eqq26wk6) (eqq26wk1->eqq26wk2) (eqq26wk2->eqq26wk3) (eqq26wk3->eqq26wk4) (eqq26wk4-
>eqq26wk5) (eqq26wk5->eqq26wk6) (dummy_icontrol frequentSW rsepre30cat freqweighXtx
rsepre30Xtx->adj_eqq16wk2 adj_eqq16wk3 adj_eqq16wk4 adj_eqq16wk5 adj_eqq16wk6)
(dummy_icontrol  mean_edewc mean_rsepre30 mean_edewc2 mean_rsepre302 -> eqq26wk2
eqq26wk3 eqq26wk4 eqq26wk5 eqq26wk6), cov(e.adj_eqq16wk2*e.adj_eqq16wk3)
cov(e.adj_eqq16wk3*e.adj_eqq16wk4) cov(e.adj_eqq16wk4*e.adj_eqq16wk5)
cov(e.adj_eqq16wk5*e.adj_eqq16wk6) cov(e.eqq26wk2*e.eqq26wk3) cov(e.eqq26wk3*e.eqq26wk4)
cov(e.eqq26wk4*e.eqq26wk5) cov(e.eqq26wk5*e.eqq26wk6) cov(_Ex) method(mlmv) allmissing
estat teffects
mat bi = r(indirect)
mat bt = r(total)
return scalar biwk2  = bi[1,13]
```

```
return scalar biwk3  = bi[1,34]
return scalar biwk4  = bi[1,58]
return scalar biwk5  = bi[1,85]
return scalar biwk6  = bi[1,115]
end
bootstrap r(biwk2) r(biwk3) r(biwk4) r(biwk5) r(biwk6), reps(1000) seed(1552470950): indireff_FW
estat bootstrap, bc
program drop indireff_FW
*standardised
sem (adj_eqq16wk1->adj_eqq16wk2 eqq26wk2) (adj_eqq16wk2->adj_eqq16wk3 eqq26wk2
eqq26wk3) (adj_eqq16wk3->adj_eqq16wk4 eqq26wk3 eqq26wk4) (adj_eqq16wk4->adj_eqq16wk5
eqq26wk4 eqq26wk5) (adj_eqq16wk5->adj_eqq16wk6 eqq26wk5 eqq26wk6) (adj_eqq16wk6->
eqq26wk6) (eqq26wk1->eqq26wk2) (eqq26wk2->eqq26wk3) (eqq26wk3->eqq26wk4) (eqq26wk4-
>eqq26wk5) (eqq26wk5->eqq26wk6) (dummy_icontrol frequentSW rsepre30cat freqweighXtx
rsepre30Xtx->adj_eqq16wk2 adj_eqq16wk3 adj_eqq16wk4 adj_eqq16wk5 adj_eqq16wk6)
(dummy_icontrol  mean_edewc mean_rsepre30 mean_edewc2 mean_rsepre302 -> eqq26wk2
eqq26wk3 eqq26wk4 eqq26wk5 eqq26wk6), cov(e.adj_eqq16wk2*e.adj_eqq16wk3)
cov(e.adj_eqq16wk3*e.adj_eqq16wk4) cov(e.adj_eqq16wk4*e.adj_eqq16wk5)
cov(e.adj_eqq16wk5*e.adj_eqq16wk6) cov(e.eqq26wk2*e.eqq26wk3) cov(e.eqq26wk3*e.eqq26wk4)
cov(e.eqq26wk4*e.eqq26wk5) cov(e.eqq26wk5*e.eqq26wk6) cov(_Ex) vce(bootstrap, reps(1000)
seed(1552470950)) method(mlmv) allmissing standardized
```

Appendix 31: Output for study 2b. Autoregressive structural equation model of the relationships between the intervention (CBT-E), the purported mediator (frequency of weighing; FW) and the outcome (concern about weight; CW)

Model Components	B	Boot. SE	Bias-corrected 95% CI of B		β
			Lower	**Upper**	
FW1→FW2	0.69	0.16	0.32	1.01	0.81
TX→FW2	0.61	0.73	-0.79	2.08	0.07
Freq weigher→FW2	1.48	0.95	-0.26	3.61	0.16
RSE group→FW2	-0.84	0.70	-2.26	0.57	-0.09
Freq weigher*TX→FW2	-3.73	1.14	-6.32	-1.65	-0.32
RSE group*TX→FW2	0.34	1.01	-1.78	2.21	0.03
FW2→CW2	0.04	0.04	-0.04	0.12	0.08
FW1→CW2	-0.00	0.04	-0.09	0.07	-0.01
WW1→CW2	0.70	0.11	0.47	0.91	0.60
TX→CW2	0.08	0.20	-0.34	0.48	0.02
WC→CW2	0.27	0.14	0.01	0.55	0.20
RSE→CW2	-0.01	0.03	-0.06	0.05	-0.01
WC(quad)→CW2	-0.05	0.05	-0.15	0.05	-0.07
RSE(quad)→CW2	0.01	0.01	0.00	0.02	0.14
FW2→FW3	0.37	0.20	0.09	0.72	0.55
TX→FW3	-0.35	0.53	-1.72	0.52	-0.06
Freq Weigher→FW3	1.92	1.15	-0.22	3.94	0.30
RSE group→FW3	-0.09	0.61	-1.65	0.82	-0.02
Freq weigher*TX→FW3	-2.36	1.14	-4.53	-0.21	-0.31
RSE group*TX→FW3	0.76	0.71	-0.32	2.57	0.11
FW2→CW3	-0.06	0.06	-0.14	0.08	-0.16
WW2→CW3	0.97	0.14	0.68	1.25	0.98
FW3→CW3	0.06	0.08	-0.09	0.20	0.10
TX→CW3	0.01	0.25	-0.48	0.47	0.00
WC→CW3	-0.08	0.13	-0.37	0.15	-0.06
RSE→CW3	0.03	0.04	-0.04	0.10	0.06
WC(quad)→CW3	0.01	0.05	-0.08	0.11	0.01
RSE(quad)→CW3	-0.01	0.01	-0.03	0.00	-0.14
FW3→FW4	1.22	0.23	0.76	1.72	1.04
TX→FW4	-0.79	0.46	-1.99	-0.08	-0.11
Freq Weigher→FW4	-0.87	0.84	-2.94	0.55	-0.12
RSE group→FW4	-1.38	0.60	-2.86	-0.38	-0.19
Freq weigher*TX→FW4	0.72	0.84	-0.71	2.75	0.08
RSE group*TX→FW4	1.26	0.63	0.15	2.72	0.16
FW3→CW4	-0.15	0.06	-0.28	-0.03	-0.24
WW3→CW4	0.92	0.12	0.70	1.17	0.93
FW3→CW4	0.16	0.05	0.07	0.26	0.31
TX→CW4	0.29	0.23	-0.16	0.72	0.08
WC→CW4	0.09	0.10	-0.12	0.28	0.07
RSE→CW4	-0.00	0.03	-0.06	0.06	-0.01
WC(quad)→CW4	0.04	0.05	-0.04	0.13	0.07
RSE(quad)→CW4	0.00	0.01	-0.01	0.02	0.05

Model Components	B	Boot. SE	Bias-corrected 95% CI of B		β
			Lower	**Upper**	
FW4→FW5	0.87	0.10	0.67	1.02	0.91
TX→FW5	-0.56	0.79	-2.58	0.50	-0.08
Freq Weigher→FW5	-0.30	0.68	-1.83	0.83	-0.04
RSE group→FW5	-0.18	0.66	-1.75	0.84	-0.03
Freq weigher*TX→FW5	0.36	0.69	-0.83	1.95	0.04
RSE group*TX→FW5	0.31	0.67	-0.75	1.96	0.04
FW4→CW5	-0.05	0.07	-0.21	0.05	-0.09
WW4→CW5	0.70	0.16	0.38	0.99	0.68
FW5→CW5	0.07	0.07	-0.07	0.23	0.13
TX→CW5	-0.19	0.25	-0.71	0.27	-0.05
WC→CW5	0.17	0.15	-0.13	0.49	0.12
RSE→CW5	-0.01	0.04	-0.08	0.06	-0.02
WC(quad)→CW5	-0.09	0.05	-0.19	0.00	-0.13
RSE(quad)→CW5	0.01	0.01	0.00	0.02	0.11
FW5→FW6	0.70	0.14	0.42	1.02	0.87
TX→FW6	-0.08	0.40	-0.95	0.60	-0.01
Freq Weigher→FW6	0.10	0.46	-0.77	1.04	0.02
RSE group→FW6	-0.50	0.41	-1.33	0.25	-0.09
Freq weigher*TX→FW6	-0.13	0.45	-1.06	0.74	-0.02
RSE group*TX→FW6	0.38	0.43	-0.39	1.21	0.06
FW5→CW6	-0.06	0.09	-0.26	0.12	-0.10
WW5→CW6	1.21	0.19	0.92	1.71	1.15
FW6→CW6	0.11	0.10	-0.09	0.31	0.15
TX→CW6	0.28	0.29	-0.22	0.91	0.07
WC→CW6	-0.13	0.16	-0.50	0.15	-0.09
RSE→CW6	-0.00	0.05	-0.10	0.08	-0.01
WC(quad)→CW6	0.11	0.07	-0.02	0.26	0.15
RSE(quad)→CW6	-0.01	0.01	-0.03	0.01	-0.09
Var(e.FW2)	6.64	1.74	4.05	11.47	0.31
Var(e.CW2)	1.26	0.24	0.89	1.94	0.34
Var(e.FW3)	3.40	1.08	1.71	6.35	0.36
Var(e.CW3)	1.31	0.21	1.02	2.04	0.37
Var(e.FW4)	2.99	1.03	1.40	5.57	0.23
Var(e.CW4)	1.22	0.22	0.91	2.07	0.35
Var(e.FW5)	2.91	1.80	0.63	7.92	0.25
Var(e.CW5)	1.48	0.21	1.15	1.99	0.40
Var(e.FW6)	0.87	0.55	0.47	1.96	0.11
Var(e.CW6)	1.67	0.53	1.06	3.57	0.41
Cov(e.FW2*e.FW3)	1.89	1.17	-0.13	4.10	0.40
Cov(e.CW2*e.CW3)	-0.38	0.19	-0.86	-0.08	-0.29
Cov(e.FW3*e.FW4)	-1.47	0.57	-2.68	-0.50	-0.46
Cov(e.CW3*e.CW4)	-0.54	0.23	-1.20	-0.21	-0.43
Cov(e.FW4*e.FW5)	-0.53	0.35	-1.31	0.08	-0.18
Cov(e.CW4*e.CW5)	-0.21	0.19	-0.64	0.10	-0.15
Cov(e.FW5*e.FW6)	0.57	0.76	-0.24	2.68	0.36
Cov(e.CW5*e.CW6)	-0.90	0.39	-2.00	-0.34	-0.57

Model Components	B	Boot. SE	Bias-corrected 95% CI of B		β
			Lower	**Upper**	
Cov(FW1*WW1)	1.95	0.62	0.96	3.67	0.22
Cov(FW1*TX)	-0.10	0.25	-0.52	0.47	-0.04
Cov(FW1*Freq Weigher)	1.44	0.23	1.07	2.01	0.55
Cov(FW1*RSE group)	-0.43	0.26	-1.01	0.03	-0.16
Cov(FW1*Freq Weigher*TX)	0.69	0.30	0.23	1.45	0.32
Cov(FW1*RSE group*TX)	-0.35	0.19	-0.74	0.02	-0.14
Cov(FW1*WC)	2.00	0.78	0.83	4.18	0.26
Cov(FW1*self-esteem)	-3.19	1.29	-6.08	-1.00	-0.16
Cov(FW1*WC(quad))	-0.70	1.11	-2.61	1.76	-0.05
Cov(FW1*self-esteem(quad))	-9.56	5.82	-23.97	0.20	-0.09
Cov(CW1*TX)	0.05	0.07	-0.08	0.20	0.06
Cov(WW1*Freq Weigher)	0.15	0.07	0.03	0.30	0.19
Cov(CW1*RSE group)	-0.09	0.07	-0.23	0.04	-0.11
Cov(CW1*Freq Weigher*TX)	0.07	0.06	-0.04	0.18	0.11
Cov(CW1*RSE group*TX)	-0.01	0.06	-0.14	0.12	-0.01
Cov(CW1*WC)	1.53	0.27	1.01	2.06	0.67
Cov(CW1*self-esteem)	-0.76	0.46	-1.66	0.07	-0.13
Cov(CW1*WC(quad))	-2.20	0.66	-3.72	-1.03	-0.48
Cov(CW1*self-esteem(quad))	0.39	2.12	-3.84	4.66	0.01
Cov(TX*Freq Weigher)	-0.01	0.02	-0.05	0.03	-0.05
Cov(TX*RSE group)	0.03	0.02	-0.02	0.07	0.11
Cov(TX*Freq Weigher*TX)	0.10	0.02	0.07	0.13	0.51
Cov(TX*RSE group*TX)	0.16	0.02	0.13	0.19	0.68
Cov(TX*WC)	0.07	0.06	-0.04	0.20	0.11
Cov(TX*self-esteem)	0.23	0.16	-0.08	0.53	0.13
Cov(TX*WC(quad))	-0.13	0.12	-0.39	0.10	-0.09
Cov(TX*self-esteem(quad))	0.04	0.86	-1.83	1.55	0.00
Cov(Freq Weigher*RSE group)	-0.04	0.02	-0.08	0.01	-0.15
Cov(Freq Weigher*Freq Weigher*TX)	0.11	0.02	0.08	0.14	0.58
Cov(Freq Weighe *RSE group*TX)	-0.03	0.02	-0.07	0.01	-0.13
Cov(Freq Weigher *WC)	0.12	0.06	0.01	0.24	0.18
Cov(Freq Weigher*self-esteem)	-0.26	0.16	-0.57	0.04	-0.15
Cov(Freq Weigher*WC(quad))	-0.29	0.11	-0.53	-0.11	-0.21
Cov(Freq Weigher *self-esteem(quad))	-0.11	0.89	-1.68	1.84	-0.01
Cov(RSE group*Freq Weigher*TX)	-0.01	0.02	-0.05	0.02	-0.07
Cov(RSE group*RSE group*TX)	0.13	0.01	0.11	0.16	0.58
Cov(RSE group*WC)	-0.07	0.06	-0.20	0.04	-0.10
Cov(RSE group*self-esteem)	1.37	0.10	1.20	1.56	0.77
Cov(RSE group*WC(quad))	0.06	0.12	-0.17	0.32	0.04
Cov(RSE group*self-esteem(quad))	0.27	0.75	-1.11	1.75	0.03
Cov(Freq Weigher*TX*RSE group*TX)	0.04	0.02	0.01	0.08	0.22
Cov(Freq Weigher*TX*WC)	0.07	0.04	-0.01	0.15	0.12
Cov(Freq Weigher*TX*self-esteem)	-0.04	0.13	-0.27	0.24	-0.03
Cov(Freq Weigher*TX*WC(quad))	-0.15	0.07	-0.30	-0.01	-0.13
Cov(Freq Weigher*TX*self-esteem(quad))	-0.03	0.78	-1.24	1.93	-0.00

Model Components	B	Boot. SE	Bias-corrected 95% CI of B		β
			Lower	**Upper**	
Cov(RSE group*TX*WC)	-0.02	0.06	-0.15	0.08	-0.03
Cov(RSE group*TX*self-esteem)	0.79	0.13	0.56	1.07	0.48
Cov(RSE group*TX*WC(quad))	-0.09	0.11	-0.29	0.12	-0.07
Cov(RSE group*TX *self-esteem(quad))	0.36	0.86	-1.20	2.26	0.04
Cov(WC*self-esteem)	-0.90	0.51	-2.02	-0.05	-0.18
Cov(WC*WC(quad))	-2.27	0.67	-3.77	-1.10	-0.57
Cov(WC*self-esteem(quad))	-3.67	3.31	-11.29	1.22	-0.13
Cov(self-esteem*WC(quad))	1.71	1.13	-0.42	3.97	0.17
Cov(self-esteem*self-esteem(quad))	31.43	14.67	6.10	63.19	0.45
Cov(WC(quad)*self-esteem(quad))	6.81	7.99	-5.05	26.18	0.13
Successful bootstraps	991/1000				

Notes. B=unstandardized coefficient. SE=standard error. β=standardised coefficient. CI=confidence interval. FW# = frequency of weighing where # indicates intervention week. CW# = concern about weight where # indicates intervention week. Freq weigher=frequent weigher status (frequent weigher vs. non-frequent weigher), RSE group=self-esteem group (high vs. low self-esteem), TX=treatment allocation, WC = weight concern at baseline, RSE=self-esteem at baseline, (quad)=quadratic trend in the covariate.

Statistical Output and Stata Code for Study 3: Frequency of Shape Checking as a Mediator of the Effect of Enhanced Cognitive Behavioural Therapy on Concern about Shape

Appendix 32: Stata code for study 3a multilevel models and their coefficient plots

```
**********CODE FOR FINAL MULTILEVEL MODELS**********
/*
Notes on variables.
- trialno = patient ID
- checking = frequency of shape checking
- center_checking = frequency of shape checking (centred around sample mean at baseline)
- eqq27 = concern about shape
- dummy_txweek = treatment week, where zero indicates baseline (start of treatment)
- dummy_txweek2 = treatment week as a quadratic trend
- time = shape checking intervention intervention week
- dummy_icontrol = treatment allocation where 0 is IPT-ED, and 1 is CBT-E
- mean_eddur = eating disorder duration at baseline centred at the sample mean
- mean_edewc = weight concern at baseline centred at the sample mean
- mean_age = age at baseline centred at the sample mean
- mean_bdi = depression at baseline centred at the sample mean
- mean_imp_shape = importance of shape at baseline centred at the sample mean
*/

****STEP 1- IMPACT OF THE INTERVENTION ON THE PURPORTED MEDIATOR****
*run final model
bootstrap, cluster(trialno) reps(1000) seed(642283768): xtmixed checking dummy_txweek
dummy_txweek2 time  dummy_icontrol dummy_icontrol#c.dummy_txweek
dummy_icontrol#c.dummy_txweek2  mean_eddur mean_edewc || trialno: dummy_txweek time,
reml cov(unstr) var
estat ic
estat icc
estat boot, bc
*create summary table of model
eststo est1
estout est1 , cells("b(fmt(2)) se p(fmt(3)) ci_bc[1] ci_bc[2]") /*
*/ modelwidth(9) style(fixed) stats(aic N_clust N N_reps)
*create coefficient plot of key variables
coefplot est1, keep(dummy_txweek dummy_txweek2 time 1.dummy_icontrol#?.dummy_txweek
1.dummy_icontrol#?.dummy_txweek2) xline(0) ci(ci_bc) ciopts(recast(rcap)) mlabel format(%9.2f)
mlabposition(12) mlabgap(*1) name(SCMLMcoeffinterval, replace) xtitle("{bf:Effect on Frequency of
Shape Checking}", height(5) ) title("{bf:1. Impact of Intervention on Purported Mediator}",
size(medlarge) color(black) span) coeflabel(dummy_txweek = "Treatment week"  dummy_txweek2 =
"Treatment week (quadratic trend)" 1.dummy_icontrol#?.dummy_txweek = "CBT-E weekly change"
1.dummy_icontrol#?.dummy_txweek2 = "CBT-E weekly change (quadratic trend)" time =
"Intervention week", wrap(40)) order (dummy_txweek dummy_txweek2
1.dummy_icontrol#?.dummy_txweek  1.dummy_icontrol#?.dummy_txweek2 time)
headings(mean_edewc = "{bf:Baseline Predictors}"  dummy_txweek  = "{bf:Impact of Treatment
```

```
Week}" 1.dummy_icontrol#?.dummy_txweek  = "{bf:Impact of Treatment Allocation}"  time  =
"{bf:Impact of Shape Checking Intervention}")

****STEP 2 - IMPACT OF THE INTERVENTION ON THE OUTCOME****
*run final model
bootstrap, cluster(trialno) reps(1000) seed(642283768): xtmixed eqq27 dummy_txweek
dummy_txweek2 time  mean_age  mean_edewc mean_bdi  mean_imp_shape
c.mean_edewc#c.dummy_txweek c.mean_bdi#c.time c.mean_imp_shape#c.time  || trialno:
dummy_txweek time, reml cov(unstr) var
estat ic
estat icc
estat boot, bc
*create summary table of model
eststo est2
estout est2, cells("b(fmt(2)) se p(fmt(3)) ci_bc[1] ci_bc[2]") modelwidth(9) style(fixed) stats(aic
N_clust N N_reps)
*create coefficient plot of key variables
coefplot est2, keep(dummy_txweek dummy_txweek2 time) xline(0) ci(ci_bc) ciopts(recast(rcap))
mlabel format(%9.2f) mlabposition(12) mlabgap(*1) name(WSMLMcoeffinterval_sc, replace)
xtitle("{bf:Effect on Concern about shape}", height(5)) title("{bf:2. Impact of Intervention on
Outcome}", size(medlarge) color(black) span) coeflabel(  dummy_txweek = "Treatment week"
dummy_txweek2 = "Treatment week (quadratic trend)" time = "Intervention week", wrap(40)) order
( dummy_txweek dummy_txweek2 1.dummy_icontrol#?.dummy_txweek
1.dummy_icontrol#?.dummy_txweek2 time) headings(mean_edewc = "{bf:Baseline Predictors}"
dummy_txweek  = "{bf:Impact of Treatment Week}" time  = "{bf:Impact of Shape Checking
Intervention}")

****STEP 3 - IMPACT OF THE PURPORTED MEDIATOR ON THE OUTCOME****
*run final model
bootstrap, cluster(trialno) reps(1000) seed(642283768): xtmixed eqq27 dummy_txweek
dummy_txweek2 center_checking  mean_edewc || trialno: dummy_txweek time , reml cov(unstr)
var
estat bootstrap, bc
estat icc
estat ic
*create summary table of model
eststo est3
estout est3, cells("b(fmt(2)) se p(fmt(3)) ci_bc[1] ci_bc[2]") modelwidth(9) style(fixed) stats(aic
N_clust N N_reps)
*create coefficient plot of key variables
coefplot est3, keep(dummy_txweek dummy_txweek2 time center_checking
c.center_checking#c.time ) xline(0) ci(ci_bc) ciopts(recast(rcap)) mlabel format(%9.2f)
mlabposition(12) mlabgap(*1) name(SCWSMLMcoeffinterval_sc, replace) xtitle("{bf:Effect on
Concern about shape}", height(5) ) title("{bf:3. Impact of Purported Mediator on Outcome}",
size(medlarge) color(black) span) coeflabel(  dummy_txweek = "Treatment week"  dummy_txweek2 =
"Treatment week (quadratic trend)" center_checking= "Frequency of shape checking" , wrap(40))
order ( dummy_txweek dummy_txweek2 center_checking ?.center_checking#c.time)
headings(dummy_txweek  = "{bf:Impact of Treatment Week}" center_checking  = "{bf:Purported
Mediator Effects}")

****STEP 4 - IMPACT OF THE INTERVENTION ON THE OUTCOME WHEN INCLUDING THE PURPORTED
MEDIATOR IN THE MODEL (CONTEMPORANEOUS RELATIONSHIP)****
*run final model
bootstrap, cluster(trialno) reps(1000) seed(642283768): xtmixed eqq27 dummy_txweek
dummy_txweek2 time   mean_edewc mean_bdi  mean_imp_shape  c.mean_bdi#c.time
```

```
c.mean_imp_shape#c.time   center_checking c.center_checking#c.time || trialno: dummy_txweek
time, reml cov(unstr) var
estat ic
estat icc
estat boot, bc
*create summary table of model
eststo est4
estout est4, cells("b(fmt(2)) se p(fmt(3)) ci_bc[1] ci_bc[2]") modelwidth(9) style(fixed) stats(aic
N_clust N N_reps)
*create coefficient plot of key variables
coefplot est4, keep(dummy_txweek dummy_txweek2 time center_checking
c.center_checking#c.time ) xline(0) ci(ci_bc) ciopts(recast(rcap)) mlabel format(%9.2f)
mlabposition(12) mlabgap(*1) name(SCmedcontempcoeffinterval_sc, replace) xtitle("{bf:Effect on
Concern about shape}", height(5) ) title("{bf:4. Impact of Intervention on Outcome with Purported
Mediator}", size(medlarge) color(black) span) coeflabel(  dummy_txweek = "Treatment week"
dummy_txweek2 = "Treatment week (quadratic trend)" time = "Intervention week" center_checking=
"Frequency of shape checking" ?.center_checking#c.time= "In intervention group", wrap(40)) order (
dummy_txweek dummy_txweek2 time center_checking ?.center_checking#c.time)
headings(mean_edewc = "{bf:Baseline Predictors}"  dummy_txweek  = "{bf:Impact of Treatment
Week}" time  = "{bf:Impact of Shape Checking Intervention}" center_checking  = "{bf:Purported
Mediator Effects}")

*COMBINE COEFPLOTS INTO ONE GRAPH
graph combine SCMLMcoeffinterval WSMLMcoeffinterval_sc SCWSMLMcoeffinterval_sc
SCmedcontempcoeffinterval_sc, name(combSC_MLM, replace) xsize(8)
```

Appendix 33: Output for study 3a, step 1. Multi-level model of the impact of the shape checking intervention on the purported mediator (frequency of shape checking)

Model Components	Variable Details	B	Boot. SE	p	Bias-corrected 95% CI of B	
					Lower	**Upper**
Baseline Predictors	Weight concern	0.15	0.05	.001	0.06	0.23
	Self-esteem	-0.05	0.02	.000	-0.08	-0.02
	Eating disorder duration	-0.02	0.01	.007	-0.03	-0.01
	Treatment allocation (CBT-E vs. IPT-ED)	0.15	0.17	.373	-0.17	0.49
Impact of treatment week	Treatment week (linear trend)	0.01	0.02	.767	-0.03	0.05
	Treatment week (quadratic)	0.00	0.00	.267	0.00	0.00
Impact of treatment allocation	CBT-E condition*Treatment week (linear trend)	-0.05	0.03	.122	-0.11	0.01
	CBT-E condition*Treatment week (quadratic trend)	0.01	0.00	.023	0.00	0.01
Impact of shape checking intervention	Shape checking intervention week (linear trend)	-0.18	0.05	.000	-0.26	-0.08
Impact of baseline predictors during treatment/intervention	Self-esteem*Intervention week (linear trend)	0.01	0.00	.018	0.00	0.01
Constant	Patients in IPT-ED with average weight concern and eating disorder duration at baseline	3.11	0.11	.000	2.87	3.32
Random effects (log values)	SD (treatment week)	-2.38	0.08	.000	-2.69	-2.32
	SD (time)	-1.95	0.12	.000	-2.31	-1.82
	SD (constant)	0.02	0.05	.638	-0.13	0.09
	Correlation (treatment week*intervention week)	-0.79	0.11	.000	-0.90	-0.48
	Correlation (treatment week*constant)	-0.53	0.08	.000	-0.65	-0.31
	Correlation (intervention week*constant)	-0.12	0.11	.267	-0.56	0.00
	SD (residual)	-0.51	0.04	.000	-0.54	-0.44
Statistics	AIC	4154.31				
	Number of patients	125				
	Number of observations	1849				
	Number of bootstraps	1000/1000				

Notes. B=unstandardized coefficient, *p*-value based on 95% confidence interval, CI=confidence interval.

Appendix 34: Output for study 3a, step 2. Multi-level model of the impact of the shape checking intervention on the outcome (concern about shape)

Model Components	Variable Details	B	Boot. SE	p	Bias-corrected 95% CI of B	
					Lower	**Upper**
Baseline Predictors	Weight concern	0.55	0.07	.000	0.38	0.68
	Weight concern (quadratic trend)	-0.09	0.03	.008	-0.15	-0.03
	Age	0.02	0.01	.005	0.01	0.04
	Depression	0.01	0.01	.100	0.00	0.02
	Importance of shape	-0.20	0.07	.005	-0.33	-0.05
Impact of treatment week	Treatment week (linear trend)	-0.13	0.02	.000	-0.17	-0.08
	Treatment week (quadratic)	0.00	0.00	.007	0.00	0.01
Impact of baseline predictors during treatment	Weight concern*Treatment week (linear trend)	-0.02	0.01	.027	-0.03	0.00
	Depression*Shape checking intervention week (linear trend)	-0.01	0.00	.022	-0.01	0.00
	Importance of shape*Shape checking intervention week (linear trend)	0.04	0.02	.008	0.02	0.08
Impact of shape checking intervention	Shape checking intervention week (linear trend)	-0.09	0.03	.002	-0.14	-0.03
Constant	Patients in the control group with average weight concern, age, depression and importance of shape at baseline	5.25	0.10	.000	5.06	5.44
Random effects (log values)	SD (treatment week)	-2.06	0.11	.000	-2.31	-2.00
	SD (time)	-1.56	0.11	.000	-1.94	-1.47
	SD (constant)	-0.12	0.10	.247	-0.49	0.02
	Correlation (treatment week*intervention week)	-0.91	0.10	.000	-0.99	-0.65
	Correlation (treatment week*constant)	-0.01	0.13	.934	-0.11	0.33
	Correlation (intervention week*constant)	-0.20	0.13	.105	-0.52	-0.10
	SD (residual)	-0.06	0.04	.151	-0.09	0.00
Statistics	AIC	5751.51				
	Number of patients	127				
	Number of observations	1854				
	Number of bootstraps	1000/1000				

Notes. B=unstandardized coefficient, *p*-value based on 95% confidence interval, CI=confidence interval.

Appendix 35: Output for study 3a, step 3. Multi-level model of impact of the purported mediator (frequency of shape checking) at time (T) on the outcome (concern about shape) at time (T)

Model Components	Variable Details	B	Boot. SE	p	Bias-corrected 95% CI of B	
					Lower	**Upper**
Baseline	Weight concern	0.50	0.06	.000	0.38	0.63
Growth trends	Treatment week (linear trend)	-0.12	0.02	.000	-0.17	-0.08
	Treatment week (quadratic)	0.00	0.00	.008	0.00	0.01
Purported mediator effects	Frequency of Shape checking	0.42	0.05	.000	0.32	0.53
Constant	Patients with average weight concern, age, depression, importance of shape and frequency of shape checking at baseline	5.09	0.07	.000	4.93	5.22
Random effects (log values)	SD (treatment week)	-2.16	0.12	.000	-2.45	-2.11
	SD (time)	-1.76	0.10	.000	-1.98	-1.70
	SD (constant)	-0.09	0.09	.344	-0.38	0.02
	Correlation (treatment week*intervention week)	-0.72	0.10	.000	-0.73	-0.61
	Correlation (treatment week*constant)	0.06	0.13	.652	-0.02	0.33
	Correlation (intervention week*constant)	-0.25	0.11	.022	-0.55	-0.23
	SD (residual)	-0.10	0.04	.016	-0.13	-0.03
Model statistics	AIC	5635.03				
	Number of patients	129				
	Number of observations	1880				
	Number of bootstraps	1000/1000				

Notes. B=unstandardized coefficient, *p*-value based on 95% confidence interval, CI=confidence interval.

Appendix 36: Output for study 3a, step 4. Multi-level model of impact of the shape checking intervention on the outcome (concern about shape) when accounting for the purported mediator (frequency of shape checking) within the same treatment week

Model Components	Variable Details	B	Boot SE	p	Bias-corrected 95%CI of B	
					Lower	**Upper**
Baseline Predictors	Weight concern	0.52	0.07	.000	0.37	0.64
	Weight concern (quadratic trend)	-0.09	0.03	.007	-0.15	-0.03
	Depression	0.01	0.01	.061	0.00	0.02
	Importance of shape	-0.19	0.07	.009	-0.32	-0.04
Impact of treatment week	Treatment week (linear trend)	-0.13	0.02	.000	-0.18	-0.08
	Treatment week (quadratic)	0.00	0.00	.002	0.00	0.01
	Depression*Shape checking intervention week (linear trend)	0.00	0.00	.039	-0.01	0.00
	Importance of shape*Shape checking intervention week (linear trend)	0.04	0.02	.018	0.02	0.08
Impact of shape checking intervention	Shape checking intervention week (linear trend)	0.01	0.04	.780	-0.07	0.07
Impact of purported mediator	Frequency of shape checking	0.35	0.05	.000	0.27	0.47
	Frequency of shape checking*intervention week	0.05	0.02	.004	0.01	0.08
Constant	Patients in the control group with average weight concern, age, depression and importance of shape at baseline	5.28	0.10	.000	5.08	5.48
Random effects (log values)	SD (treatment week)	-2.13	0.11	.000	-2.38	-2.10
	SD (time)	-1.69	0.11	.000	-2.07	-1.63
	SD (constant)	-0.11	0.10	.302	-0.45	0.04
	Correlation (treatment week*intervention week)	-0.78	0.13	.000	-0.90	-0.41
	Correlation (treatment week*constant)	0.01	0.13	.927	-0.09	0.35
	Correlation (intervention week*constant)	-0.16	0.13	.218	-0.52	-0.05
	SD (residual)	-0.10	0.04	.012	-0.13	-0.04
Statistics	AIC	5591.11				
	Number of patients	127				
	Number of observations	1850				
	Number of bootstraps	1000/1000				

Notes. B=unstandardized coefficient, *p*-value based on 95% confidence interval, CI=confidence interval.

Appendix 37: Stata code for study 3b autoregressive structural equation model

```
*******CODE FOR FINAL AUTOREGRESSIVE STRUCTURAL EQUATION MODEL*******
/*
Notes on variables.
- trialno = patient ID
- checking_time99 = frequency of shape checking at baseline
- checking_time0 = frequency of shape checking at the start of the intervention
- checking_time100 = frequency of shape checking 1 week into the intervention
- checking_time200 = frequency of shape checking 2 week into the intervention
- checking_time300 = frequency of shape checking 3 week into the intervention
- checking_time400 = frequency of shape checking 4 week into the intervention
- checking_time500 = frequency of shape checking 5 week into the intervention
- checking_time600 = frequency of shape checking 6 week into the intervention
- eqq27_time99 = concern about shape at baseline
- eqq27_time0 = concern about shape at the start of the intervention
- eqq27_time100 = concern about shape 1 week into the intervention
- eqq27_time200 = concern about shape 2 week into the intervention
- eqq27_time300 = concern about shape 3 week into the intervention
- eqq27_time400 = concern about shape 4 week into the intervention
- eqq27_time500 = concern about shape 5 week into the intervention
- eqq27_time600 = concern about shape 6 week into the intervention
- dummy_icontrol = treatment allocation where 0 is IPT-ED, and 1 is CBT-E
- receive_txs = intervention status where 0 is the 'control' group and 1 is the 'intervention' group
- mean_eddur = eating disorder duration at baseline centred at the sample mean
- mean_edewc = level of  weight concern at baseline centred at the sample mean
- mean_age = age at baseline centred at the sample mean
- mean_bdi = level of depression at baseline centred at the sample mean
- mean_imp_shape = degree of importance of shape at baseline centred at the sample mean
*/

*unstandardised model
sem (checking_time0->checking_time100 eqq27_time100) (checking_time100->checking_time200
eqq27_time100 eqq27_time200)  (checking_time200->checking_time300 eqq27_time200
eqq27_time300) (checking_time300->checking_time400 eqq27_time300 eqq27_time400)
(checking_time400->checking_time500 eqq27_time400 eqq27_time500) (checking_time500-
>checking_time600 eqq27_time500 eqq27_time600) (checking_time600->eqq27_time600)
(eqq27_time0->eqq27_time100) (eqq27_time100->eqq27_time200) (eqq27_time200-
>eqq27_time300) (eqq27_time300->eqq27_time400)  (eqq27_time400->eqq27_time500)
(eqq27_time500->eqq27_time600) (receive_txs-> checking_time100 checking_time200
checking_time300 checking_time400 checking_time500 checking_time600  eqq27_time100
eqq27_time200 eqq27_time300 eqq27_time400 eqq27_time500 eqq27_time600)  (mean_edewc
mean_eddur checking_time99->  checking_time100 checking_time200 checking_time300
checking_time400 checking_time500 checking_time600)  (mean_edewc mean_age mean_bdi
mean_imp_shape eqq27_time99->  eqq27_time100 eqq27_time200 eqq27_time300 eqq27_time400
eqq27_time500 eqq27_time600) ,  cov(e.checking_time100*e.checking_time200)
cov(e.checking_time200*e.checking_time300) cov(e.checking_time300*e.checking_time400)
cov(e.checking_time400*e.checking_time500) cov(e.checking_time500*e.checking_time600)
*/ cov(e.eqq27_time100*e.eqq27_time200) cov(e.eqq27_time200*e.eqq27_time300)
cov(e.eqq27_time300*e.eqq27_time400)  cov(e.eqq27_time400*e.eqq27_time500)
cov(e.eqq27_time500*e.eqq27_time600) cov(_Ex) vce(bootstrap, reps(1000) seed(196906640))
method(mlmv) allmissing
estat boot, bc
*fit indices
estat gof, stats(all)
```

```
*test stability of path coefficients of checking
test(_b[checking_time100:checking_time0]==_b[checking_time200:checking_time100]==_b[checking
_time300:checking_time200]==_b[checking_time400:checking_time300] /*
*/ ==_b[checking_time500:checking_time400]==_b[checking_time600:checking_time500] )
*test stability of path coefficients of concern about shape
test(_b[eqq27_time100:eqq27_time0]==_b[eqq27_time200:eqq27_time100]==_b[eqq27_time300:e
qq27_time200]==_b[eqq27_time400:eqq27_time300]/*
*/ ==_b[eqq27_time500:eqq27_time400]==_b[eqq27_time600:eqq27_time500] )
*get table of direct, indirect and total effects of treatment, set matrices & return scalars for model to
run bootstrap teffects
estat teffects, standardized
mat list r(direct)
mat list r(indirect)
mat list r(total)
program indireff, rclass
set more off
sem (checking_time0->checking_time100 eqq27_time100) (checking_time100->checking_time200
eqq27_time100 eqq27_time200)  (checking_time200->checking_time300 eqq27_time200
eqq27_time300) (checking_time300->checking_time400 eqq27_time300 eqq27_time400)
(checking_time400->checking_time500 eqq27_time400 eqq27_time500) (checking_time500-
>checking_time600 eqq27_time500 eqq27_time600) (checking_time600->eqq27_time600)
(eqq27_time0->eqq27_time100) (eqq27_time100->eqq27_time200) (eqq27_time200-
>eqq27_time300) (eqq27_time300->eqq27_time400) (eqq27_time400->eqq27_time500)
(eqq27_time500->eqq27_time600) (receive_txs-> checking_time100 checking_time200
checking_time300 checking_time400 checking_time500 checking_time600  eqq27_time100
eqq27_time200 eqq27_time300 eqq27_time400 eqq27_time500 eqq27_time600)  (mean_edewc
mean_eddur checking_time99->  checking_time100 checking_time200 checking_time300
checking_time400 checking_time500 checking_time600)  (mean_edewc mean_age mean_bdi
mean_imp_shape eqq27_time99->  eqq27_time100 eqq27_time200 eqq27_time300 eqq27_time400
eqq27_time500 eqq27_time600),  cov(e.checking_time100*e.checking_time200)
cov(e.checking_time200*e.checking_time300) cov(e.checking_time300*e.checking_time400)
cov(e.checking_time400*e.checking_time500) cov(e.checking_time500*e.checking_time600)
cov(e.eqq27_time100*e.eqq27_time200) cov(e.eqq27_time200*e.eqq27_time300)
cov(e.eqq27_time300*e.eqq27_time400)  cov(e.eqq27_time400*e.eqq27_time500)
cov(e.eqq27_time500*e.eqq27_time600) cov(_Ex) method(mlmv) allmissing
estat teffects
mat bi = r(indirect)
mat bt = r(total)
return scalar biwk1  = bi[1,9]
return scalar biwk2  = bi[1,28]
return scalar biwk3  = bi[1,50]
return scalar biwk4  = bi[1,75]
return scalar biwk5  = bi[1,103]
return scalar biwk6  = bi[1,134]
return scalar btwk1  = bt[1,9]
return scalar btwk2  = bt[1,28]
return scalar btwk3  = bt[1,50]
return scalar btwk4  = bt[1,75]
return scalar btwk5  = bt[1,103]
return scalar btwk6  = bt[1,134]
end
bootstrap r(biwk1) r(biwk2) r(biwk3) r(biwk4) r(biwk5) r(biwk6) r(btwk1) r(btwk2) r(btwk3) r(btwk4)
r(btwk5) r(btwk6), reps(1000) seed(196906640): indireff
estat bootstrap, bc
program drop indireff
```

```
*standardised
sem (checking_time0->checking_time100 eqq27_time100) (checking_time100->checking_time200
eqq27_time100 eqq27_time200)  (checking_time200->checking_time300 eqq27_time200
eqq27_time300) (checking_time300->checking_time400 eqq27_time300 eqq27_time400)
(checking_time400->checking_time500 eqq27_time400 eqq27_time500) (checking_time500-
>checking_time600 eqq27_time500 eqq27_time600) (checking_time600->eqq27_time600)
(eqq27_time0->eqq27_time100) (eqq27_time100->eqq27_time200) (eqq27_time200-
>eqq27_time300) (eqq27_time300->eqq27_time400)  (eqq27_time400->eqq27_time500)
(eqq27_time500->eqq27_time600) (receive_txs-> checking_time100 checking_time200
checking_time300 checking_time400 checking_time500 checking_time600  eqq27_time100
eqq27_time200 eqq27_time300 eqq27_time400 eqq27_time500 eqq27_time600)  (mean_edewc
mean_eddur checking_time99->  checking_time100 checking_time200 checking_time300
checking_time400 checking_time500 checking_time600)  (mean_edewc mean_age mean_bdi
mean_imp_shape eqq27_time99->  eqq27_time100 eqq27_time200 eqq27_time300 eqq27_time400
eqq27_time500 eqq27_time600),  cov(e.checking_time100*e.checking_time200)
cov(e.checking_time200*e.checking_time300) cov(e.checking_time300*e.checking_time400)
cov(e.checking_time400*e.checking_time500) cov(e.checking_time500*e.checking_time600)
cov(e.eqq27_time100*e.eqq27_time200) cov(e.eqq27_time200*e.eqq27_time300)
cov(e.eqq27_time300*e.eqq27_time400)  cov(e.eqq27_time400*e.eqq27_time500)
cov(e.eqq27_time500*e.eqq27_time600) cov(_Ex) standardized  vce(bootstrap, reps(1000)
seed(196906640)) method(mlmv) allmissing
```

Appendix 38: Output for study 3b. Autoregressive partial mediation model of the relationships between the shape checking intervention, frequency of shape checking at the previous week and within the same week, and concern about shape at time (T) using full information maximum likelihood estimation for missing data

Model Components	B	Boot. SE	Bias-corrected 95% CI of B		β
			Lower	**Upper**	
FC0→FC1	0.64	0.08	0.49	0.80	0.66
SCINT→FC1	0.40	0.13	0.14	0.65	0.17
WC→FC1	0.00	0.05	-0.10	0.10	-0.01
EDDUR→FC1	0.00	0.01	-0.02	0.01	-0.01
SELF-ESTEEM→FC1	0.00	0.02	-0.04	0.04	-0.01
FCBASELINE→FC1	0.21	0.09	0.03	0.38	0.23
FC1→CS1	0.34	0.15	0.09	0.69	0.22
FC0→CS1	-0.05	0.15	-0.37	0.22	-0.04
CS0→CS1	0.73	0.08	0.56	0.85	0.70
SCINT→CS1	-0.01	0.24	-0.50	0.43	0.00
WC→CS1	-0.13	0.12	-0.38	0.08	-0.10
WC(QUAD)→CS1	-0.02	0.05	-0.12	0.07	-0.04
AGE→CS1	-0.01	0.01	-0.04	0.01	-0.03
DEP→CS1	-0.01	0.01	-0.03	0.01	-0.06
IMP SHAPE→CS1	0.15	0.08	-0.01	0.31	0.13
FCBASELINE→CS1	0.12	0.10	-0.07	0.31	0.08
FC1→FC2	0.93	0.12	0.71	1.21	0.98
SCINT→FC2	-0.17	0.17	-0.53	0.16	-0.08
WC→FC2	0.01	0.06	-0.09	0.11	0.01
EDDUR→FC2	0.00	0.01	-0.02	0.01	-0.02
SELF-ESTEEM→FC2	0.00	0.02	-0.04	0.04	-0.01
FCBASELINE→FC2	-0.19	0.11	-0.42	0.01	-0.21
FC1→CS2	-0.06	0.15	-0.38	0.21	-0.04
CS1→CS2	0.79	0.12	0.57	1.06	0.78
FC2→CS2	0.35	0.14	0.06	0.60	0.21
SCINT→CS2	0.32	0.25	-0.17	0.84	0.09
WC→CS2	0.14	0.14	-0.16	0.39	0.11
WC(QUAD)→CS2	0.03	0.06	-0.08	0.14	0.04
AGE→CS2	0.01	0.01	-0.01	0.04	0.05
DEP→CS2	-0.02	0.01	-0.04	0.00	-0.14
IMP SHAPE→CS2	0.00	0.12	-0.25	0.24	0.00
FCBASELINE→CS2	-0.03	0.12	-0.28	0.21	-0.02
FC2→FC3	0.85	0.11	0.62	1.05	0.83
SCINT→FC3	-0.49	0.16	-0.82	-0.18	-0.22
WC→FC3	0.05	0.06	-0.05	0.17	0.07
EDDUR→FC3	0.00	0.01	-0.02	0.02	0.02
SELF-ESTEEM→FC3	0.02	0.03	-0.03	0.08	0.07
FCBASELINE→FC3	0.09	0.07	-0.04	0.25	0.10
FC2→CS3	-0.46	0.17	-0.81	-0.15	-0.28
CS2→CS3	0.98	0.19	0.62	1.39	0.98
FC3→CS3	0.37	0.15	0.06	0.66	0.23

Model Components	B	Boot. SE	Bias-corrected 95% CI of B		β
			Lower	**Upper**	
SCINT→CS3	-0.15	0.30	-0.76	0.47	-0.04
WC→CS3	0.15	0.20	-0.24	0.55	0.12
WC(QUAD)→CS3	0.00	0.06	-0.12	0.12	0.00
AGE→CS3	0.01	0.02	-0.03	0.04	0.02
DEP→CS3	0.01	0.02	-0.01	0.04	0.08
IMP SHAPE→CS3	-0.12	0.16	-0.43	0.19	-0.11
FCBASELINE→CS3	0.03	0.22	-0.39	0.45	0.02
FC3→FC4	0.91	0.13	0.67	1.19	0.90
SCINT→FC4	-0.24	0.18	-0.59	0.13	-0.10
WC→FC4	0.03	0.05	-0.08	0.14	0.03
EDDUR→FC4	0.01	0.01	-0.01	0.03	0.08
SELF-ESTEEM→FC4	0.02	0.02	-0.03	0.06	0.06
FCBASELINE→FC4	0.03	0.10	-0.19	0.22	0.03
FC3→CS4	-0.78	0.21	-1.21	-0.41	-0.48
CS3→CS4	1.15	0.15	0.90	1.47	1.13
FC4→CS4	0.64	0.17	0.31	0.97	0.39
SCINT→CS4	-0.32	0.30	-0.95	0.28	-0.09
WC→CS4	-0.33	0.21	-0.77	0.06	-0.25
WC(QUAD)→CS4	-0.01	0.07	-0.13	0.13	-0.01
AGE→CS4	-0.03	0.02	-0.07	0.01	-0.11
DEP→CS4	0.03	0.02	-0.01	0.06	0.16
IMP SHAPE→CS4	0.22	0.16	-0.07	0.56	0.20
FCBASELINE→CS4	-0.13	0.16	-0.48	0.16	-0.09
FC4→FC5	0.93	0.15	0.62	1.19	0.92
SCINT→FC5	0.05	0.21	-0.35	0.49	0.02
WC→FC5	0.05	0.07	-0.09	0.16	0.06
EDDUR→FC5	-0.01	0.01	-0.03	0.02	-0.04
SELF-ESTEEM→FC5	0.00	0.02	-0.05	0.05	0.01
FCBASELINE→FC5	-0.01	0.10	-0.22	0.17	-0.01
FC4→CS5	-0.39	0.20	-0.92	-0.06	-0.24
CS4→CS5	0.87	0.12	0.68	1.15	0.86
FC5→CS5	0.50	0.20	0.12	0.91	0.31
SCINT→CS5	-0.02	0.29	-0.60	0.49	-0.01
WC→CS5	0.02	0.16	-0.29	0.33	0.01
WC(QUAD)→CS5	-0.02	0.06	-0.15	0.10	-0.04
AGE→CS5	0.04	0.02	0.01	0.09	0.17
DEP→CS5	-0.02	0.02	-0.05	0.01	-0.12
IMP SHAPE→CS5	-0.15	0.12	-0.41	0.07	-0.13
FCBASELINE→CS5	0.15	0.12	-0.07	0.37	0.10
FC5→FC6	0.93	0.12	0.74	1.22	0.97
SCINT→FC6	-0.45	0.17	-0.79	-0.12	-0.20
WC→FC6	0.00	0.06	-0.13	0.12	0.00
EDDUR→FC6	-0.01	0.01	-0.03	0.01	-0.05
SELF-ESTEEM→FC6	0.03	0.02	-0.01	0.08	0.11
FCBASELINE→FC6	-0.02	0.10	-0.22	0.16	-0.02
FC5→CS6	-0.54	0.22	-0.97	-0.08	-0.33

Model Components	B	Boot. SE	Bias-corrected 95% CI of B		β
			Lower	**Upper**	
CS5→CS6	1.14	0.14	0.91	1.45	1.14
FC6→CS6	0.58	0.19	0.25	0.95	0.34
SCINT→CS6	0.21	0.35	-0.49	0.87	0.06
WC→CS6	0.32	0.18	-0.03	0.68	0.25
WC(QUAD)→CS6	0.05	0.07	-0.10	0.18	0.07
AGE→CS6	-0.04	0.03	-0.10	0.01	-0.15
DEP→CS6	0.00	0.02	-0.04	0.03	-0.01
IMP SHAPE→CS6	-0.15	0.15	-0.47	0.15	-0.13
FCBASELINE→CS6	-0.27	0.12	-0.52	-0.06	-0.19
Var(e.FC1)	0.41	0.06	0.33	0.57	0.31
Var(e.CS1)	1.10	0.20	0.80	1.57	0.35
Var(e.FC2)	0.61	0.10	0.45	0.90	0.51
Var(e.CS2)	1.30	0.17	1.09	1.91	0.41
Var(e.FC3)	0.58	0.08	0.46	0.86	0.47
Var(e.CS3)	1.67	0.34	1.21	2.83	0.52
Var(e.FC4)	0.71	0.13	0.51	1.05	0.56
Var(e.CS4)	1.88	0.43	1.37	4.00	0.56
Var(e.FC5)	0.67	0.14	0.45	1.09	0.52
Var(e.CS5)	1.44	0.22	1.16	2.44	0.42
Var(e.FC6)	0.52	0.15	0.35	1.14	0.44
Var(e.CS6)	1.31	0.37	0.84	2.87	0.39
Cov(e.FC1*e.FC2)	-0.16	0.06	-0.32	-0.07	-0.32
Cov(e.CS1*e.CS2)	-0.33	0.16	-0.79	-0.07	-0.28
Cov(e.FC2*e.FC3)	-0.18	0.09	-0.40	-0.03	-0.30
Cov(e.CS2*e.CS3)	-0.81	0.25	-1.51	-0.43	-0.55
Cov(e.FC3*e.FC4)	-0.29	0.09	-0.61	-0.17	-0.46
Cov(e.CS3*e.CS4)	-1.00	0.28	-1.96	-0.62	-0.56
Cov(e.FC4*e.FC5)	-0.26	0.13	-0.60	-0.07	-0.38
Cov(e.CS4*e.CS5)	-0.52	0.18	-1.08	-0.28	-0.32
Cov(e.FC5*e.FC6)	-0.30	0.12	-0.76	-0.14	-0.51
Cov(e.CS5*e.CS6)	-0.83	0.26	-1.61	-0.44	-0.61
Cov(FC0* CS0)	0.88	0.17	0.54	1.23	0.44
Cov(FC0*SCINT)	0.01	0.05	-0.09	0.12	0.02
Cov(FC0*WC)	0.43	0.14	0.17	0.71	0.26
Cov(FC0*EDDUR)	-1.47	0.91	-3.36	0.21	-0.14
Cov(FC0*SELF-ESTEEM)	-0.86	0.38	-1.69	-0.19	-0.20
Cov(FC0*FCBASELINE)	0.80	0.13	0.56	1.06	0.55
Cov(FC0*WC(QUAD))	-0.76	0.28	-1.38	-0.26	-0.23
Cov(FC0*AGE)	-0.87	0.80	-2.44	0.73	-0.10
Cov(FC0*DEP)	2.86	1.08	0.76	4.95	0.22
Cov(FC0*IMP SHAPE)	0.25	0.16	-0.09	0.56	0.13
Cov(FC0*CSBASELINE)	0.26	0.14	0.00	0.52	0.18
Cov(CS0*SCINT)	-0.01	0.08	-0.17	0.13	-0.02
Cov(CS0*WC)	0.84	0.25	0.35	1.34	0.35
Cov(CS0*EDDUR)	2.84	1.32	0.37	5.58	0.19
Cov(CS0*SELF-ESTEEM)	-1.05	0.53	-2.11	-0.01	-0.17

Model Components	B	Boot. SE	Bias-corrected 95% CI of B		β
			Lower	**Upper**	
Cov(CS0*FCBASELINE)	0.28	0.20	-0.10	0.64	0.13
Cov(CS0*WC(QUAD))	-1.58	0.52	-2.62	-0.60	-0.33
Cov(CS0*AGE)	2.78	1.17	0.34	5.00	0.21
Cov(CS0*DEP)	3.83	1.66	0.72	7.04	0.21
Cov(CS0*IMP SHAPE)	0.42	0.26	-0.06	0.97	0.15
Cov(CS0*FCBASELINE	0.92	0.27	0.45	1.56	0.43
Cov(SCINT*WC)	0.01	0.06	-0.10	0.12	0.01
Cov(SCINT*EDDUR)	-0.46	0.36	-1.17	0.23	-0.11
Cov(SCINT*SELF-ESTEEM)	0.18	0.16	-0.15	0.47	0.10
Cov(SCINT*FCBASELINE)	0.03	0.05	-0.07	0.14	0.05
Cov(SCINT*WC(QUAD))	-0.14	0.11	-0.37	0.07	-0.10
Cov(SCINT *AGE)	-0.21	0.32	-0.82	0.41	-0.06
Cov(SCINT *DEP)	-0.61	0.46	-1.47	0.35	-0.12
Cov(SCINT*IMP SHAPE)	0.08	0.06	-0.05	0.21	0.10
Cov(SCINT*CSBASELINE	-0.06	0.06	-0.19	0.04	-0.10
Cov(WC*EDDUR)	1.17	1.15	-1.21	3.31	0.10
Cov(WC*SELF-ESTEEM)	-0.94	0.53	-2.01	-0.04	-0.19
Cov(WC*FCBASELINE)	0.47	0.16	0.15	0.79	0.27
Cov(WC*WC(QUAD))	-2.27	0.68	-3.73	-1.09	-0.57
Cov(WC*AGE)	0.45	1.10	-1.96	2.42	0.04
Cov(WC*DEP)	4.72	1.47	1.95	7.52	0.31
Cov(WC*IMP SHAPE)	1.67	0.30	1.07	2.25	0.73
Cov(WC*CSBASELINE)	0.95	0.20	0.59	1.40	0.54
Cov(WC*SELF-ESTEEM)	1.59	2.31	-2.71	6.41	0.05
Cov(EDDUR*FCBASELINE)	-2.81	1.03	-4.92	-0.89	-0.26
Cov(EDDUR*WC(QUAD))	1.30	2.40	-2.25	7.75	0.05
Cov(EDDUR*AGE)	58.08	12.63	36.45	86.77	0.88
Cov(EDDUR*DEP)	12.67	8.67	-4.30	30.44	0.14
Cov(EDDUR*IMP SHAPE)	1.37	1.31	-1.28	3.88	0.10
Cov(EDDUR*CSBASELINE)	1.41	0.75	0.05	3.10	0.13
Cov(SELF-ESTEEM *FCBASELINE)	-1.14	0.39	-1.90	-0.35	-0.25
Cov(SELF-ESTEEM *WC(QUAD))	1.76	1.14	-0.21	4.15	0.17
Cov(SELF-ESTEEM *AGE)	2.03	2.08	-2.12	5.91	0.07
Cov(SELF-ESTEEM *DEP)	-11.98	3.27	-18.83	-6.19	-0.31
Cov(SELF-ESTEEM *IMP SHAPE)	-0.81	0.71	-2.32	0.46	-0.14
Cov(SELF-ESTEEM *CSBASELINE)	-0.47	0.39	-1.20	0.35	-0.10
Cov(FCBASELINE*WC(QUAD))	-0.91	0.31	-1.58	-0.34	-0.26
Cov(FCBASELINE*AGE)	-2.37	0.91	-4.16	-0.49	-0.25
Cov(FCBASELINE*DEP)	1.03	1.14	-1.22	3.15	0.08
Cov(FCBASELINE*IMP SHAPE)	0.37	0.17	0.04	0.72	0.19
Cov(FCBASELINE *CSBASELINE)	0.49	0.13	0.23	0.77	0.31
Cov(WC(QUAD)*AGE)	2.25	2.41	-1.22	8.61	0.10
Cov(WC(QUAD)*DEP)	-2.79	3.14	-9.51	2.56	-0.09
Cov(WC(QUAD)*IMP SHAPE)	-2.68	0.79	-4.35	-1.21	-0.59
Cov(WC(QUAD)*CSBASELINE)	-1.29	0.41	-2.12	-0.53	-0.37
Cov(AGE*DEP)	13.27	7.79	-3.68	28.67	0.16

Model Components	B	Boot. SE	Bias-corrected 95% CI of B		β
			Lower	**Upper**	
Cov(AGE*IMP SHAPE)	0.84	1.19	-2.00	2.89	0.07
Cov(AGE*CSBASELINE	1.24	0.67	-0.01	2.68	0.13
Cov(DEP*IMP SHAPE)	3.88	1.76	0.79	7.63	0.22
Cov(DEP*CCBASELINE)	2.84	1.15	0.79	5.28	0.21
Cov(IMP SHAPE *CSBASELINE	0.68	0.21	0.33	1.16	0.34
Successful bootstraps	986/1000				

Notes. B=unstandardized coefficient. SE=standard error. β=standardised coefficient. CI=confidence interval. FC# = frequency of shape checking, where # indicates intervention week. SC# = concern about shape, where # indicates intervention week. SCINT = intervention status (shape checking intervention vs. control). WC = weight concern at baseline. EDDUR = eating disorder duration at baseline. FCBASELINE = frequency of shape checking at baseline. AGE = age at baseline. DEP = depression at baseline. IMP SHAPE= importance of shape at baseline. CSBASELINE = concern about shape at baseline.

Statistical Output and Stata Code for Study 4: Interpersonal Problem Severity as a Mediator of the Effect of Interpersonal Psychotherapy for Eating Disorders on Eating Disorder Psychopathology

Appendix 39: Stata code for study 4a multilevel models and their coefficient plots

```
**********CODE FOR FINAL MULTILEVEL MODELS**********
/*
Notes on variables.
- trialno = patient ID
- ipi_prob1sev = interpersonal problem severity
- center_ipi_prob1sev = interpersonal problem severity (centred around sample mean at baseline)
- glede = eating disorder psychopathology
- dummy_txweek = treatment week, where zero indicates baseline (start of treatment)
- dummy_txweek2 = treatment week as a quadratic trend
- dummy_ccontrol = treatment allocation where 0 is CBT-E, and 1 is IPT-ED
- mean_eddur = eating disorder duration at baseline centred at the sample mean
- mean_bdi = depression at baseline centred at the sample mean
- mean_rsepre30 = self-esteem at baseline centred at the sample mean
*/

****STEP 1- IMPACT OF THE INTERVENTION ON THE PURPORTED MEDIATOR****
*run final model
bootstrap, cluster(trialno) reps(1000) seed(729123051): xtmixed ipi_prob1sev time time2
dummy_ccontrol dummy_ccontrol#c.time mean_eddur dummy_diag mean_bdi mean_rsepre30
c.mean_rsepre30#c.mean_rsepre30 || trialno: , reml cov(unstr) var
estat ic
estat icc
estat bootstrap, bc
*create summary table of model
eststo est1
estout est1, cells("b(fmt(2)) se p(fmt(3)) ci_bc[1] ci_bc[2]") modelwidth(9) style(fixed) stats(aic
N_clust N N_reps)
*create coefficient plot of key variables
coefplot est1, keep(time time2  1.dummy_ccontrol#?.time) xline(0) ci(ci_bc) ciopts(recast(rcap))
mlabel format(%9.2f) mlabposition(12) mlabgap(*1) name(IPsevMLMcoeffinterval_norandoms,
replace) xtitle("{bf:Effect on Interpersonal Problem Severity}", height(5) ) title("{bf:1. Impact of
Intervention on Purported Mediator}", size(medlarge) color(black) span) coeflabel(time = "Time"
time2 = "Time (quadratic trend)" 1.dummy_ccontrol#?.time = "IPT-ED condition (linear change)",
wrap(40)) order (time time2 1.dummy_ccontrol#?.time) headings( time  = "{bf:Impact of Time}"
1.dummy_ccontrol#?.time  = "{bf:Impact of Treatment Allocation}")
saving(IPsevMLMcoeffinterval_norandoms, replace)

****STEP 2 - IMPACT OF THE INTERVENTION ON THE OUTCOME****
*run final model
```

```
bootstrap, cluster(trialno) reps(1000) seed(729123051): xtmixed glede time time2 dummy_ccontrol dummy_ccontrol#c.time dummy_ccontrol#c.time2 c.mean_eddur c.mean_eddur#c.time c.mean_bdi || trialno: , reml cov(unstr) var
estat ic
estat icc
estat bootstrap, bc
*create summary table of model
eststo est2
estout est2, cells("b(fmt(2)) se p(fmt(3)) ci_bc[1] ci_bc[2]") modelwidth(9) style(fixed) stats(aic N_clust N N_reps)
*create coefficient plot of key variables
coefplot est2, keep(time time2  1.dummy_ccontrol#?.time 1.dummy_ccontrol#?.time2) xline(0) ci(ci_bc) ciopts(recast(rcap)) mlabel format(%9.2f) mlabposition(12) mlabgap(*1) name(gledeMLMcoeffinterval_norandom, replace) xtitle("{bf:Effect on Eating Disorder Psychopathology}", height(5) ) title("{bf:2. Impact of Intervention on Outcome}", size(medlarge) color(black) span) coeflabel(time = "Time"  time2 = "Time (quadratic trend)" 1.dummy_ccontrol#?.time = "IPT-ED condition (linear change)" 1.dummy_ccontrol#?.time2="IPT-ED (quadratic trend)", wrap(40)) order (time time2 1.dummy_ccontrol#?.time 1.dummy_ccontrol#?.time2) headings(time  = "{bf:Impact of Time}"  1.dummy_ccontrol#?.time  = "{bf:Impact of Treatment Allocation}") saving(gledeMLMcoeffinterval_norandom, replace)

****STEP 3 - IMPACT OF THE PURPORTED MEDIATOR ON THE OUTCOME****
*run final model
bootstrap, cluster(trialno) reps(1000) seed(729123051): xtmixed glede time time2 center_ipi_prob1sev c.mean_eddur c.mean_bdi  || trialno: , reml cov(unstr) var
estat ic
estat icc
estat bootstrap, bc
*create summary table of model
eststo est3
estout est3, cells("b(fmt(2)) se p(fmt(3)) ci_bc[1] ci_bc[2]") modelwidth(9) style(fixed) stats(aic N_clust N N_reps)
*create coefficient plot of key variables
coefplot est3, keep(time time2  center_ipi_prob1sev) xline(0) ci(ci_bc) ciopts(recast(rcap)) mlabel format(%9.2f) mlabposition(12) mlabgap(*1) name(IPTmotxcovMLMcoeffint_norandom, replace) xtitle("{bf:Effect on Eating Disorder Psychopathology}", height(5)) title("{bf:3. Impact of Purported Mediator on Outcome}", size(medlarge) color(black) span) coeflabel(center_ipi_prob1sev = "IP Problem Severity (T)" time = "Time" time2 = "Time (quadratic trend)", wrap(40)) order (time time2 center_ipi_prob1sev) headings(time  = "{bf:Impact of Time}" center_ipi_prob1sev  = "{bf:Impact of Purported Mediator}") saving(IPTmotxcovMLMcoeffint_norandom, replace)

****STEP 4 - IMPACT OF THE INTERVENTION ON THE OUTCOME WHEN INCLUDING THE PURPORTED MEDIATOR IN THE MODEL (CONTEMPORANEOUS RELATIONSHIP)****
*run final model
bootstrap, cluster(trialno) reps(10) seed(729123051): xtmixed glede time time2 dummy_ccontrol dummy_ccontrol#c.time dummy_ccontrol#c.time2  c.mean_bdi center_ipi_prob1sev || trialno: , reml cov(unstr) var
estat ic
estat icc
estat bootstrap, bc
*create summary table of model
eststo est4
estout est4, cells("b(fmt(2)) se p(fmt(3)) ci_bc[1] ci_bc[2]") modelwidth(9) style(fixed) stats(aic N_clust N N_reps)
*create coefficient plot of key variables
```

```
coefplot est4, keep(time time2 1.dummy_ccontrol#?.time 1.dummy_ccontrol#?.time2
center_ipi_prob1sev) xline(0) ci(ci_bc) ciopts(recast(rcap)) mlabel format(%9.2f) mlabposition(12)
mlabgap(*1) name(IPTmedMLMcoeffinterval_norandom, replace) xtitle("{bf:Effect on Eating Disorder
Psychopathology}", height(5)) title("{bf:4. Impact of Intervention on Outcome with Purported
Mediator}", size(medlarge) color(black) span) coeflabel(time = "Time" time2 = "Time (quadratic
trend)" 1.dummy_ccontrol#?.time = "IPT-ED condition (linear change)"
1.dummy_ccontrol#?.time2="IPT-ED (quadratic trend)"  center_ipi_prob1sev = "IP Problem Severity
(T)", wrap(40)) order (time time2 1.dummy_ccontrol#?.time 1.dummy_ccontrol#?.time2
center_ipi_prob1sev) headings(mean_bdi = "{bf:Baseline Predictors}"  time  = "{bf:Impact of Time}"
1.dummy_ccontrol#?.time  = "{bf:Impact of Treatment Allocation}"  center_ipi_prob1sev  =
"{bf:Impact of Purported Mediator}") saving(IPTmedMLMcoeffinterval_norandom, replace)

*COMBINE COEFPLOTS INTO ONE GRAPH
graph combine IPsevMLMcoeffinterval_norandoms gledeMLMcoeffinterval_norandom
IPTmotxcovMLMcoeffint_norandom IPTmedMLMcoeffinterval_norandom,
name(combIPTstudy_MLM_sametxoutcovs, replace) xsize(8)
```

Appendix 40: Output for study 4a, step 1. Multi-level model of the impact of IPT-ED on the purported mediator (interpersonal problem severity)

Model Components	Variable Details	B	Boot. SE	p	Bias-corrected 95% CI of B	
					Lower	**Upper**
Baseline Predictors	Depression	0.02	0.01	.007	0.00	0.03
	Self-esteem	0.02	0.02	.425	-0.03	0.06
	Self-esteem (quadratic trend)	-0.01	0.00	.013	-0.02	-0.00
	Eating disorder duration	0.02	0.01	.054	0.00	0.04
	Eating disorder diagnosis (BN vs. EDNOS)	0.35	0.17	.037	0.02	0.68
	IPT-ED condition	0.18	0.20	.363	-0.23	0.55
Impact of Time	Time (linear trend)	-2.02	0.21	.000	-2.41	-1.61
	Time (quadratic)	0.40	0.06	.000	0.27	0.52
Impact of treatment allocation	IPT-ED condition*Time (linear trend)	-0.01	0.13	.933	-0.27	0.28
Constant	Patients with average weight concern and eating disorder duration in CBT-E at baseline	3.69	0.16	.000	3.40	4.00
Random effects (log values)	SD (constant)	-0.27	0.08	.001	-0.48	-0.21
	SD (residual)	0.23	0.05	.000	0.21	0.30
Statistics	AIC	1457.12				
	Number of patients	119				
	Number of observations	396				
	Number of bootstraps	1000/1000				

Notes. B=unstandardized coefficient, *p*-value based on 95% confidence interval, CI=confidence interval. BN=bulimia nervosa. EDNOS=eating disorder not otherwise specified.

Appendix 41: Output for study 4a, step 2. Multi-level model of the impact of IPT-ED on the outcome (eating disorder psychopathology)

Model Components	Variable Details	B	Boot. SE	p	Bias-corrected 95% CI of B	
					Lower	**Upper**
Baseline Predictors	Eating disorder duration	0.00	0.01	.715	-0.01	0.02
	Depression	0.03	0.01	.000	0.01	0.04
	IPT-ED condition	-0.02	0.16	.919	-0.33	0.29
Impact of Time	Time (linear trend)	-1.92	0.20	.000	-2.30	-1.51
	Time (quadratic)	0.44	0.06	.000	0.32	0.55
Impact of baseline predictors during treatment	Eating disorder duration*Time (linear trend)	0.01	0.00	.009	0.00	0.02
Impact of treatment allocation	IPT-ED condition*Time (linear trend)	0.59	0.28	.034	0.07	1.12
	IPT-ED condition*Time (quadratic trend)	-0.16	0.09	.060	-0.34	-0.01
Constant	Patients with average weight concern and eating disorder duration in CBT-E at baseline	3.49	0.11	.000	3.24	3.69
Random effects (log values)	SD (constant)	-0.30	0.07	.000	-0.51	-0.24
	SD (residual)	-0.11	0.05	.024	-0.12	-0.05
Statistics	AIC	1391.90				
	Number of patients	128				
	Number of observations	450				
	Number of bootstraps	1000/1000				

Notes. B=unstandardized coefficient, *p*-value based on 95% confidence interval, CI=confidence interval.

Appendix 42: Output for study 4a, step 3. Multi-level model of the impact of the purported mediator (interpersonal problem severity) on the outcome (eating disorder psychopathology)

Model Components	Variable Details	B	Boot. SE	p	Bias-corrected 95% CI of B	
					Lower	**Upper**
Baseline Predictors	Eating disorder duration	0.02	0.01	.023	0.00	0.03
	Depression	0.03	0.01	.000	0.02	0.04
Impact of Time	Time (linear trend)	-1.34	0.15	.000	-1.64	-1.04
	Time (quadratic)	0.31	0.05	.000	0.23	0.41
Purported Mediator Effects	Interpersonal problem severity (centred at baseline)	0.18	0.04	.000	0.11	0.27
Constant	Patients with average weight concern and eating disorder duration at baseline	3.44	0.08	.000	3.29	3.58
Random effects (log values)	SD (constant)	-0.39	0.07	.000	-0.56	-0.34
	SD (residual)	-0.14	0.04	.001	-0.14	-0.09
Statistics	AIC	1200.94				
	Number of patients	122				
	Number of observations	402				
	Number of bootstraps	1000/1000				

Notes. B=unstandardized coefficient, *p*-value based on 95% confidence interval, CI=confidence interval.

Appendix 43: Output for study 4a, step 4. Multi-level model of the impact of IPT-ED on the outcome (eating disorder psychopathology) when accounting for the purported mediator (interpersonal problem severity) within the same time point

Model Components	Variable Details	B	Boot. SE	p	Bias-corrected 95% CI of B	
					Lower	**Upper**
Baseline Predictors	Depression	0.03	0.01	.000	0.02	0.04
	Treatment allocation (IPT-ED vs. CBT-E)	0.01	0.15	.940	-0.28	0.32
Impact of Time	Time (linear trend)	-1.75	0.21	.000	-2.17	-1.37
	Time (quadratic)	0.42	0.06	.000	0.31	0.55
Impact of treatment allocation	IPT-ED condition*Time (linear trend)	0.80	0.26	.002	0.31	1.32
	IPT-ED condition*Time (quadratic trend)	-0.22	0.08	.009	-0.37	-0.06
Purported Mediator Effects	Interpersonal problem severity (centred at baseline)	0.19	0.04	.000	0.12	0.27
Constant	Patients with average weight concern and eating disorder duration in CBT-E at baseline	3.43	0.11	.000	3.22	3.64
Random effects (log values)	SD (constant)	-0.42	0.07	.000	-0.66	-0.37
	SD (residual)	-0.15	0.05	.001	-0.15	-0.09
Statistics	AIC	1191.11				
	Number of patients	122				
	Number of observations	402				
	Number of bootstraps	1000/1000				

Notes. B=unstandardized coefficient, *p*-value based on 95% confidence interval, CI=confidence interval.

Appendix 44: Stata code for study 4b autoregressive structural equation model

```
*******CODE FOR FINAL AUTOREGRESSIVE STRUCTURAL EQUATION MODEL*******
/*
Notes on variables.
- trialno = patient ID
- ipi_prob1sev_time0 = interpersonal problem severity at treatment week 3
- ipi_prob1sev_time1 = interpersonal problem severity treatment week 19
- ipi_prob1sev_time2 = interpersonal problem severity at 20-week follow-up
- ipi_prob1sev_time3 = interpersonal problem severity at 40-week follow-up
- glede_time0 = eating disorder psychopathology at baseline
- glede_time1 = eating disorder psychopathology at end of treatment
- glede_time2 = eating disorder psychopathology at 20-week follow-up
- glede_time3 = eating disorder psychopathology at 40-week follow-up
- dummy_ccontrol = treatment allocation where 0 is CBT-E, and 1 is IPT-ED
- mean_eddur = eating disorder duration at baseline centred at the sample mean
- mean_bdi = depression at baseline centred at the sample mean
- mean_rsepre30 = self-esteem at baseline centred at the sample mean
- mean_rsepre30_2 = self-esteem at baseline centred at the sample mean (quadratic trend)
*/

*unstandardized model
sem (ipi_prob1sev_time0->ipi_prob1sev_time1 glede_time1) (ipi_prob1sev_time1-
>ipi_prob1sev_time2 glede_time1 glede_time2) (ipi_prob1sev_time2->ipi_prob1sev_time3
glede_time2 glede_time3) (ipi_prob1sev_time3->glede_time3) (glede_time0->glede_time1)
(glede_time1->glede_time2) (glede_time2->glede_time3) (dummy_ccontrol->ipi_prob1sev_time1
ipi_prob1sev_time2 ipi_prob1sev_time3 glede_time1 glede_time2 glede_time3) (mean_eddur
mean_bdi ->glede_time1 glede_time2 glede_time3) (mean_eddur dummy_diag mean_bdi
mean_rsepre30 mean_rsepre30_2->ipi_prob1sev_time1 ipi_prob1sev_time2 ipi_prob1sev_time3),
cov(e.ipi_prob1sev_time1*e.ipi_prob1sev_time2) cov(e.ipi_prob1sev_time2*e.ipi_prob1sev_time3)
cov(e.glede_time1*e.glede_time2) cov(e.glede_time2*e.glede_time3) cov(_Ex) vce(bootstrap,
reps(1000) seed(893578028)) method(mlmv) allmissing
estat boot, bc
*fit indices
estat gof, stats(all)
*test stability of path coefficients of interpersonal problem severity
test(_b[ipi_prob1sev_time1:ipi_prob1sev_time0]==_b[ipi_prob1sev_time2:ipi_prob1sev_time1]==_b
[ipi_prob1sev_time3:ipi_prob1sev_time2])
*test stability of path coefficients of eating disorder psychopathology
test(_b[glede_time1:glede_time0]==_b[glede_time2:glede_time1]==_b[glede_time3:glede_time2])
*teffects
estat teffects, standardized
*check matrices
mat list r(indirect)
mat list r(total)
*stop - set matrices & return scalars for model to run bootstrap teffects
program indireff, rclass
set more off
sem (ipi_prob1sev_time0->ipi_prob1sev_time1 glede_time1) (ipi_prob1sev_time1-
>ipi_prob1sev_time2 glede_time1 glede_time2) (ipi_prob1sev_time2->ipi_prob1sev_time3
glede_time2 glede_time3) (ipi_prob1sev_time3->glede_time3) (glede_time0->glede_time1)
(glede_time1->glede_time2) (glede_time2->glede_time3) (dummy_ccontrol->ipi_prob1sev_time1
ipi_prob1sev_time2 ipi_prob1sev_time3 glede_time1 glede_time2 glede_time3) (mean_eddur
mean_bdi ->glede_time1 glede_time2 glede_time3) (mean_eddur dummy_diag mean_bdi
mean_rsepre30 mean_rsepre30_2->ipi_prob1sev_time1 ipi_prob1sev_time2 ipi_prob1sev_time3),
```

```
cov(e.ipi_prob1sev_time1*e.ipi_prob1sev_time2) cov(e.ipi_prob1sev_time2*e.ipi_prob1sev_time3)
cov(e.glede_time1*e.glede_time2) cov(e.glede_time2*e.glede_time3) cov(_Ex) method(mlmv)
allmissing
estat teffects
mat bi = r(indirect)
mat bt = r(total)
return scalar bitime1  = bi[1,11]
return scalar bitime2  = bi[1,30]
return scalar bitime3  = bi[1,52]
return scalar bttime1  = bt[1,11]
return scalar bttime2  = bt[1,30]
return scalar bttime3  = bt[1,52]
end
bootstrap r(bitime1) r(bitime2) r(bitime3) r(bttime1) r(bttime2) r(bttime3), reps(1000)
seed(893578028): indireff
estat bootstrap, bc
program drop indireff
*standardised
sem (ipi_prob1sev_time0->ipi_prob1sev_time1 glede_time1) (ipi_prob1sev_time1-
>ipi_prob1sev_time2 glede_time1 glede_time2) (ipi_prob1sev_time2->ipi_prob1sev_time3
glede_time2 glede_time3) (ipi_prob1sev_time3->glede_time3) (glede_time0->glede_time1)
(glede_time1->glede_time2) (glede_time2->glede_time3) (dummy_ccontrol->ipi_prob1sev_time1
ipi_prob1sev_time2 ipi_prob1sev_time3 glede_time1 glede_time2 glede_time3) (mean_eddur
mean_bdi->glede_time1 glede_time2 glede_time3) (mean_eddur dummy_diag mean_bdi
mean_rsepre30 mean_rsepre30_2->ipi_prob1sev_time1 ipi_prob1sev_time2 ipi_prob1sev_time3),
cov(e.ipi_prob1sev_time1*e.ipi_prob1sev_time2) cov(e.ipi_prob1sev_time2*e.ipi_prob1sev_time3)
cov(e.glede_time1*e.glede_time2) cov(e.glede_time2*e.glede_time3) cov(_Ex) vce(bootstrap,
reps(1000) seed(893578028)) method(mlmv) allmissing standardized
```

Appendix 45: Output for study 4b. Autoregressive partial mediation model of the relationships IPT-ED, interpersonal problem severity at the previous time point and within the same time point, and eating disorder psychopathology using full information maximum likelihood estimation for missing data

Model Components	B	Boot. SE	Bias-corrected 95% CI of B		β
			Lower	**Upper**	
IPS0→IPS1	0.48	0.15	0.16	0.77	0.30
Treatment (IPT-ED vs. CBT-E)→IPS1	-0.22	0.36	-0.91	0.50	-0.06
ED duration→ IPS1	0.02	0.02	-0.02	0.05	0.08
Depression→IPS1	0.00	0.02	-0.03	0.04	0.01
ED diagnosis→ IPS1	0.40	0.37	-0.29	1.17	0.11
Self-esteem→ IPS1	0.01	0.05	-0.08	0.11	0.03
Self-esteem (quadratic trend)→ IPS1	-0.02	0.01	-0.03	-0.00	-0.20
IPS1→EDP1	0.25	0.07	0.10	0.37	0.34
IPS0→EDP1	-0.10	0.10	-0.28	0.11	-0.08
EDP0→EDP1	0.46	0.11	0.23	0.67	0.36
Treatment (IPT-ED vs. CBT-E)→EDP1	0.83	0.23	0.36	1.28	0.32
Eating disorder duration→EDP1	0.00	0.01	-0.02	0.03	0.02
Depression→EDP1	0.00	0.01	-0.02	0.02	0.01
IPS1→IPS2	0.31	0.34	-0.44	0.94	0.35
Treatment (IPT-ED vs. CBT-E)→IPS2	0.18	0.32	-0.43	0.82	0.06
ED duration→ IPS2	0.02	0.02	-0.01	0.06	0.10
Depression→IPS2	0.01	0.02	-0.03	0.04	0.04
ED diagnosis→ IPS2	0.42	0.39	-0.28	1.26	0.13
Self-esteem→ IPS2	0.02	0.05	-0.06	0.11	0.05
Self-esteem (quadratic trend)→ IPS2	-0.01	0.01	-0.03	0.02	-0.07
IPS1→EDP2	-0.17	0.10	-0.39	0.04	-0.23
EDP1→EDP2	0.94	0.35	0.38	1.84	0.95
IPS2→EDP2	0.18	0.09	0.00	0.35	0.22
Treatment (IPT-ED vs. CBT-E)→EDP2	-0.53	0.38	-1.46	0.06	-0.20
Eating disorder duration→EDP2	0.01	0.01	-0.01	0.04	0.08
Depression→EDP2	0.00	0.01	-0.02	0.03	0.03
IPS2→IPS3	0.73	0.64	0.24	1.50	0.76
Treatment (IPT-ED vs. CBT-E)→IPS3	-0.16	0.36	-0.88	0.42	-0.05
ED duration→ IPS3	0.02	0.02	-0.02	0.06	0.14
Depression→IPS3	0.01	0.01	-0.02	0.03	0.07
ED diagnosis→ IPS3	-0.46	0.40	-1.26	0.15	-0.15
Self-esteem→ IPS3	0.04	0.05	-0.06	0.13	0.09
Self-esteem (quadratic trend)→ IPS3	-0.01	0.01	-0.03	0.01	-0.16
IPS2→EDP3	0.02	0.20	-0.20	0.25	0.02
EDP2→EDP3	0.58	0.58	0.24	0.99	0.62
IPS3→EDP3	0.15	0.10	-0.02	0.37	0.19
Treatment (IPT-ED vs. CBT-E)→EDP3	0.27	0.19	-0.10	0.64	0.11
Eating disorder duration→EDP3	0.02	0.02	-0.00	0.04	0.12
Depression→EDP3	0.01	0.02	-0.01	0.03	0.07
var(e.IPS1)	2.77	0.28	2.42	4.01	0.85
var(e.EDP1)	1.15	0.15	0.93	1.54	0.66

Model Components	B	Boot. SE	Bias-corrected 95% CI of B		β
			Lower	**Upper**	
var(e.IPS2)	1.96	0.81	1.27	4.11	0.78
var(e.EDP2)	1.18	0.51	0.77	2.95	0.69
var(e.IPS3)	1.49	15.43	0.85	4.85	0.65
var(e.EDP3)	0.68	8.52	0.45	1.34	0.45
cov(e.IPS1*e.IPS2)	0.12	0.89	-1.45	2.08	0.05
cov(e.EDP1*e.EDP2)	-0.47	0.36	-1.37	0.08	-0.41
cov(e.IPS2*e.IPS3)	-0.51	1.24	-2.25	0.50	-0.30
cov(e.EDP2*e.EDP3)	-0.07	0.59	-0.54	0.41	-0.08
cov(IPS0*EDP0)	0.21	0.09	0.06	0.41	0.19
cov(IPS0*IPT-ED condition)	0.08	0.05	-0.01	0.18	0.14
cov(IPS0*eating disorder duration)	0.79	0.91	-1.15	2.38	0.08
cov(IPS0*depression)	4.03	1.12	2.05	6.33	0.33
cov(IPS0*ED diagnosis)	0.07	0.05	-0.03	0.17	0.13
cov(IPS0*self-esteem)	-0.38	0.41	-1.29	0.34	-0.09
cov(IPS0*self-esteem (quadratic trend))	0.30	2.52	-6.15	4.39	0.01
cov(EDP0*IPT-ED condition)	-0.02	0.05	-0.11	0.06	-0.04
cov(EDP0*eating disorder duration)	0.72	0.87	-1.27	2.24	0.08
cov(EDP0*depression)	4.32	1.12	2.36	6.62	0.39
cov(EDP0*ED diagnosis)	0.07	0.04	-0.01	0.16	0.13
cov(EDP0*self-esteem)	-0.78	0.38	-1.68	-0.15	-0.21
cov(EDP0*self-esteem (quadratic trend))	-2.85	2.85	-10.88	0.92	-0.14
cov(IPT-ED conditon*eating disorder duration)	0.75	0.36	0.08	1.48	0.17
cov(IPT-ED condition*depression)	0.39	0.47	-0.61	1.25	0.07
cov(IPT-ED condition*ED diagnosis)	0.01	0.02	-0.03	0.05	0.05
cov(IPT-ED condition*self-esteem)	-0.24	0.16	-0.54	0.09	-0.13
cov(IPT-ED condition*self-esteem (quadratic trend))	-0.06	0.89	-1.58	1.94	-0.01
cov(eating disorder duration*depression)	12.79	9.27	-5.13	30.14	0.14
cov(eating disorder duration*ED diagnosis)	0.07	0.38	-0.67	0.79	0.02
cov(eating disorder duration*self-esteem)	1.62	2.36	-2.57	6.34	0.05
cov(eating disorder duration*self-esteem (quadratic trend))	0.78	12.11	-23.31	23.15	0.00
cov(depression*ED diagnosis)	0.61	0.48	-0.32	1.62	0.11
cov(depression*self-esteem)	-11.94	3.20	-18.68	-6.03	-0.31
cov(depression*self-esteem (quadratic trend))	-32.36	18.65	-80.71	-4.27	-0.15
cov(ED diagnosis*self-esteem)	-0.14	0.17	-0.50	0.17	-0.08
cov(ED diagnosis*self-esteem (quadratic trend))	0.26	0.86	-1.41	2.02	0.03
cov(self-esteem*self-esteem (quadratic trend))	31.24	14.79	7.99	65.69	0.45
Successful Bootstraps	883/1000				

Notes. B=unstandardized coefficient. SE=standard error. β=standardised coefficient. CI=confidence interval. IPS# = interpersonal problem severity, where # indicates time point (0 = baseline, 1 = end of treatment, 2 = 20-week follow-up, 3 = 40-week follow-up. EDP# =eating disorder psychopathology, where # indicates time point. ED=eating disorder. Eating disorder duration, depression and self-esteem measured at baseline.