ANGLIA RUSKIN UNIVERSITY

FACULTY OF HEALTH, EDUCATION, MEDICINE AND SOCIAL CARE

UNDERSTANDING THERAPIST VARIABLES: AN ANALYSIS OF ONLINE COGNITIVE BEHAVIOURAL THERAPY TRANSCRIPTS

SARAH BATEUP

A thesis in partial fulfilment of the requirements of Anglia Ruskin University for the degree of Professional Doctorate

Submitted: January 2020

ACKNOWLEDGEMENTS

This is the first time in my career where I have been able to really study what cognitive behavioural therapists do with their patients. Despite managing services, teaching trainees at various universities, as well as supervising countless therapists it has never really been possible to have insight into what each therapist was doing with their patients. This is because, traditionally, therapy is delivered behind a closed door and whilst therapists are required to submit recordings of their work, these are always self-selected and represent only a tiny part of the therapist's clinical work. I must therefore thank all the therapists at Ieso Digital Health, without whom this work would not have been possible. It is an enormous privilege to study their work and, through them, I aim to use the learning to inform the work of therapists' everywhere. I am also indebted to all the patients who have had therapy on the Ieso Digital Therapy platform. Each and every one of you are helping us amplify the effect of CBT so that future patients can benefit from what we learn. This is a new way of studying psychological therapy. I feel hugely thankful for being in the right place at the right time so that I could be part of this discovery. I am extremely grateful to Barnaby Perks, Nigel Pitchford and the Board of Directors at Ieso Digital Health who enabled me to have financial support to complete this Professional Doctorate. I would like to thank all of my colleagues at Ieso Digital Health, every one of you has played an important part in supporting me to complete this work. Without the dedicated team at Ieso we would not have built a platform and a service that has delivered CBT to so many people in England. Particular thanks must go to Dr Ana Catarino who patiently taught me about statistics and listened carefully as I tried to articulate what I was doing with my statistical analyses. Ana you are an amazing teacher and colleague, thank you! I also want to thank Dr Andy Blackwell who encouraged me every step of the way. This was really invaluable particularly at those low points when I questioned why I was doing a doctorate.

My supervisory team, Dr Sarah Burch, at Anglia Ruskin University and Dr Gillian Todd, have also been there for me at those low points in the last five years. Sarah, I particularly remember your kindness and patience when I had to make the hard decision to completely change my research study at the end of 2017. I nearly gave up at this point, but you helped keep me on the right track. Gillian, your skilled support, to use self-practice and self-reflection when my critical voice has run away with me, was just what I needed. Your patience and genuine interest in my work have really helped me.

My family have also been by my side though the highs and lows. I want to thank my husband Ken, who supports me whatever I want to do, and my children who I know got totally fed up with the amount of time I was spending reading, scribbling endless notes and typing at my laptop. Finally, I want to thank my father. You are not here to know how the story ends, but I feel your love and pride. You are the man who taught me to be adventurous, to have a go… even if it felt hard, impossible or if I lost my way. Thank you.

ANGLIA RUSKIN UNIVERSITY

ABSTRACT

FACULTY OF HEALTH, EDUCATION, MEDICINE AND SOCIAL CARE

PROFESSIONAL DOCTORATE

UNDERSTANDING THERAPIST VARIABLES: AN ANALYSIS OF ONLINE COGNITIVE BEHAVIOURAL THERAPY TRANSCRIPTS

S. E. BATEUP

JANUARY 2020

OBJECTIVES
Cognitive behavioural therapy (CBT) is a psychological therapy that is widely recommended for the treatment of depression and anxiety disorders. The variance in recovery rates for CBT in England's Improving Access to Psychological Therapy (IAPT) programme has received much attention, with some services reporting recovery rates as low as 18%. One of the variables that account for clinical outcome are therapists. Without access to therapy transcripts it has been difficult to assert which therapist variables are associated with outcome. The purpose of this research is to use the therapy transcripts of 200 IAPT therapists in order to understand which therapist variables are associated with clinical outcome.

METHOD
This research used a naturalistic observational study design to understand the relationship between clinical outcomes and therapist variables. The therapy transcripts of 200 High Intensity IAPT therapists, who had provided CBT online using synchronous written communication, were rated by 6 highly experienced CBT therapists. The raters used the revised version of the Cognitive Therapy Scale to assess therapist competence and used the transcripts of 3 whole episodes of care to rate adherence to an evidence-based protocol. The data were analysed using correlation, regression and loglinear models.

RESULTS
This research found that not all therapists deliver CBT with fidelity to the model (competence) or adherence to a protocol. Where this was evident, therapist competence and therapist adherence were related to clinical outcome at the $\leq$ 0.05 level. Therapist age, gender, core profession, years of experience or method of training were not related to outcome

CONCLUSION
Process-outcome research in IAPT has, to date, been unable to access the therapy transcripts of large numbers of therapists. This is the first time that this has been possible to use therapy transcripts to understand the relationship between clinical outcome and therapist competence and adherence to an evidence-based protocol. This new way of conducting psychological therapy research provides a unique contribution to knowledge and will have a significant impact on professional practice in relation to how CBT therapists are supported to improve patient outcomes in the context of IAPT.

KEY WORDS: Cognitive behavioural therapy, IAPT, outcomes, therapist variables, competency, adherence

TABLE OF CONTENTS

Contents

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

CBT:	Cognitive Behavioural Therapy
CDS:	Clinical Decision Support
CTS-R:	Cognitive Therapy Scale-Revised
HI:	High Intensity
IAPT:	Improving Access to Psychological Therapy
IECBT:	Internet Enabled Cognitive Behavioural Therapy
NHS:	National Health Service
NICE:	National Institute for Health and Social Care Excellence
PWP:	Psychological Wellbeing Practitioner

LIST OF APPENDICES

COPYRIGHT DECLARATION

UNDERSTANDING THERAPIST VARIABLES: AN ANALYSIS OF ONLINE COGNITIVE BEHAVIOURAL THERAPY TRANSCRIPTS

SARAH BATEUP

JANUARY 2020

SCIENTIFIC PEER REVIEWED PUBLICATIONS DURING THE PERIOD OF THIS THESIS

As the Chief Clinical Officer at the company Ieso Digital Health (see www.iesohealth.com) I am part of a scientific community of clinical scientists, Artificial Intelligence scientists and clinicians. I am the senior clinician in this team. The work reported in this thesis has been conducted independently by me. In the midst of conducting this research it became clear that this work represents a new way of conducting psychological therapy research. The ability to study what therapists are doing with patients, in volume, in a real-world clinical setting has very significant implications for psychological therapy. These implications go beyond the work presented in this thesis. In parallel with this work, I have published the following papers:

Ewbank, M.P., Cummins, R., Tablan, V., **Bateup, S**., Catarino, A., Martin, A.J., Blackwell, A.D. (2019). Quantifying the association between psychotherapy content and clinical outcomes using deep learning. *JAMA Psychiatry* (see appendix item 1).

Catarino, A., **Bateup, S**., Tablan, V., Inness, K., Freer, S., Richards, A., Stott, R., Hollon, S.D., Chamberlain, S.R., Hayes, A., and Blackwell, A. (2018). Demographic and Clinical predictors of response to internet-enabled cognitive-behavioural therapy for depression and anxiety. *British Journal of Psychiatry*, 4, 411-418 (see appendix item 2).

Burch, S., Preston, C., and **Bateup, S**. (2018). True technology-enabled mental health care: Trust, agency and ageing. *mHealth* (see appendix item 3).

Burch, S., Preston, C., **Bateup, S**., and Hina, F. (2017). The use of internet-enabled cognitive behavioural therapy in the treatment of depression and anxiety amongst older people. *The International Journal of Aging and Society*, 8, 1-11 (see appendix item 4).

Bateup, S.E., Palmer, C., and Catarino, A. (2020). Using technology to understand how therapist variables are associated with clinical outcomes in IAPT.

SOURCES OF DATA USED IN THIS RESEARCH

The data that has been used in the research described in this thesis was collected as part of normal service delivery at Ieso Digital Health (see www.iesohealth.com). The company provides psychological therapy to NHS patients as part of the Improving Access to Psychological Therapy programme (IAPT). The data collected in this research falls under the auspices of the legal framework for data collection and data usage in the IAPT programme. NHS England, under section 250 of the Health and Social Care Act (2012), have directed IAPT services to collect patient data. This data set, otherwise known as the IAPT minimum data set, includes information relating to a patient's presenting problem, demographics, clinical outcomes and patient experience within the IAPT programme. The data is intended to be used to conduct clinical audits, monitor the effectiveness of the programme and the performance of individual services, national reporting analyses and research (NHS Digital 2019 a,b,c). Explicit consent from the patient is not required for the data to be used for these purposes (NHS Digital 2019 a,b,c). This is currently the case and will continue to be the case going forward (NHS Digital 2019 e).

Where IAPT services use data to support secondary uses, NHS England encourage services to require patients to opt into treatment having been made aware of how their deidentified data will be used (Clark, 2011, 2018, Gyani, Shfran, Layard and Clark, 2013). This guidance is based on Information Governance best practice at the time when the data described in this thesis was analysed. Additionally, patients may manage their preferences, in relation to how services use their data for research purposes, through the NHS National Data Opt Out Service (NHS Digital 2019d). Where a patient changes their preference and opts out, their data is no longer used for research purposes.

When registering to use the Ieso Digital Health service patients are required to 'opt in' to the service by agreeing to the services terms and conditions policy. This policy provides detailed

information in relation to how Ieso Digital Health store and use their deidentified data for service improvement, clinical audit and research. Additionally, patients are provided with a plain language guidance on how their deidentified data will be used (see appendix item 5). This enables patients to make an informed choice when opting into the service. Patients may opt out at any time and this does not affect their ongoing treatment. Where a patient opts out, their data is excluded from all research.

Additionally, it should be pointed out that the analysis of the data described in this thesis was performed as part of normal routine care. The researcher and the senior clinicians who reviewed the transcripts of therapy sessions have routine access to this data in order to assess the quality of each therapists work and to support therapists to make effective clinical decisions.

This research described in this thesis was reviewed by a Research Ethics Committee. Ethical approval for this research was granted on 17th January 2018 by the Departmental Research Ethics Panel (DREP) under the terms of Anglia Ruskin University's Research Ethics Policy, dated 8 September 2016, Version 1.7, reference FHSE-DREP-17-069. This study was also conducted in accordance with the International Conference for Harmonisation of Good Clinical Practice and the Research Governance Framework for Health and Social Care. (Health Research Authority 2016). Further discussion regarding ethics can be found in section 4.3 of this thesis.

CHAPTER ONE: AN INTRODUCTION TO THE PROFESSIONAL CONTEXT FOR THIS RESEARCH

This thesis presents research that was undertaken between February 2015 and February 2019 as part of a Professional Doctorate in Health and Social Care. Due to the professional context of this research, and the close relationship between the research and my professional role, the first and last chapters of this thesis are written in the first person as they present my personal reflections on this research. In this first chapter I discuss my rationale for my chosen research area and present an overview for this thesis. In the final chapter I present my reflections on how the process of undertaking this research has impacted on my own professional practice. The remaining chapters of this thesis are written from a scientist practitioner stance, in the third person, consistent with quantitative research.

1.1 INTRODUCTION

This research focuses on understanding the relationship between therapist variables and clinical outcome in England's Improving Access to Psychological Therapy programme (IAPT). The IAPT programme was implemented in 2008 and aims to improve access to evidenced based psychological therapies, such as cognitive behavioural therapy, to patients who present with anxiety disorders and depression (Layard and Clark, 2014). This research focuses specifically on British Association of Behavioural Cognitive Psychotherapy (BABCP) accredited therapists delivering cognitive behavioural therapy (CBT) to patients at step 3 of the IAPT programme (see Chapter 2 for an explanation of the stepped care model developed by IAPT). This thesis will present how, despite the fact that CBT has a strong empirical evidence base that supports the notion that at least 50% of all patients treated with CBT will recover, or achieve a clinically significant reliable improvement, (Layard and Clark, 2014), there is a significant variance between individual cognitive behavioural therapists and between IAPT services (Gyanni, Shafran, Layard and Clark, 2013). Additionally, IAPT has a

well-established, post-graduate, training curricula for cognitive behavioural therapy and mandatory minimum training standards for all BABCP accredited therapists (Layard and Clark 2014). Given the training and accreditation criteria are closely aligned with the evidence base for CBT, it is reasonable to assume that IAPT therapists should achieve similar results to those reported in the outcome research for CBT. Therefore, this thesis hypothesises that the variance in outcome is, in some part, directly attributed to what individual therapists are doing, or not doing with their patients.

1.2 PROFESSIONAL BACKGROUND AND CONTEXT FOR THIS RESEARCH

The research reported in this thesis not only reflects my professional background, values and beliefs, but it is rooted within my current professional role. In this section I will establish my professional background, my theoretical position, current role and how these aspects of my professional background have led to this research.

1.2.1 Professional background

I am a British Association of Behavioural and Cognitive Psychotherapy (BABCP) accredited cognitive behavioural therapist. Since qualifying as a CBT therapist in 1994 I have delivered approximately 30,000 hours of CBT to NHS patients. I have always been driven to make a difference and perhaps for this reason I set out to try and influence the work of others by going on to found two primary care services, managing CBT services and teaching on postgraduate training programmes at four universities and working as an external examiner for a post-graduate CBT training programme.

1.2.2 Cognitive behavioural therapy

I was attracted to CBT because of its structured, evidence-based approach. CBT is a psychological therapy that enables patients to develop an understanding of the predisposing

and precipitating factors that have led to them experiencing symptoms such as anxiety or depression (Beck, 2011). Furthermore, CBT places an emphasis on developing a shared understanding (between therapist and patient) of how thoughts, emotions, physiological symptoms and behaviour are interconnected and each serve to maintain the patient's presenting problem (Beck, 2011). This formulation-driven approach is fundamental in socialising the patient to the CBT model and enabling them to understand the rationale for the interventions that follow. CBT interventions are protocol-driven (Roth and Pilling, 2007, 2008) and each treatment protocol consists of a range of change mechanisms, each of which target identified cognitive or behavioural problem areas (Roth and Pilling, 2007, 2008). Therapist and patient work together with the aim of enabling the patient to learn how to reduce (or extinguish) the frequency and intensity of their symptoms using specific change mechanisms. Unlike most other forms of psychological therapy CBT places an emphasis on patients actively engaging with treatment and between-session practice of key skills and processes is routinely encouraged. Patients are supported to elicit clear and measurable, behavioural, goals (Beck, 2011). Progress towards these goals is measured at regular intervals through the use of validated, self-administered questionnaires and through regular conversational reviews of progress (Beck, 2011). The emphasis on learning as an integral part of CBT cannot be overemphasised (Beck, 2011). Patients who engage in CBT are not passive recipients of treatment but active participants who, with the support of their therapist, learn techniques and strategies that can, if practised regularly, enable them to manage their symptoms and feel better (Beck, 2011). That is not say that CBT is a panacea for all mental health conditions, and for all people, but it is currently the treatment of choice for most common health conditions (Layard and Clark, 2014) in England.

1.2.3 Scientist practitioner stance: epistemic and ontological beliefs

As I reported in 1.2.2 above, I was attracted to CBT, rather than any other sort of psychological therapy, because of its sound empirical evidence base. As a clinician it has

always been important to me to be guided by scientific principles. CBT is arguably the most evidenced based psychological therapy (Layard and Clark 2014) and its structure and theoretical framework appealed to me because I believed that these principles would give me the best chance of making a difference to patients' lives. CBT training exposed me to robust principles of quantitative research. In my work with patients I learnt to develop and test hypotheses and to reflect on each patient I worked with in order to understand why some patients got better and others did not. I sought to be a scientist practitioner, using quantitative data collection and empirical investigation (Barlow, Hayes and Nelson 1984, Milne and Paxton 1998,) to guide and drive my clinical practice. The term Scientist Practitioner, as it appertains to psychological therapists, was first used in the early 1950s at a time when evidence based psychological therapies, such as behaviour therapy, were becoming more popular (Raimy,1950). The term describes a career-long process of empirical investigation using an "experimental approach to science and psychological interventions" (Milne and Paxton 1998 p.217). Scientist Practitioners lean towards epistemic beliefs that are nomothetic rather than idiographic, preferring scientific observation rather than intuition and a scientific approach rather than a humanistic approach (Conway,1988). Ontologically, Scientist Practitioners lean towards realism believing that reality exists independent of the researcher. My training, experience and exposure to key opinion leaders in the field have led me to adopt this approach towards my own professional practice, and indeed this research. CBT has been born out of scientist experimentation since the 1950s (McHugh and Barlow, 2012). Undoubtedly, it is this tradition that has led to my preference to view my clinical work and research through a positivist lens. Whilst embracing the principles of a Scientist Practitioner I have become increasingly more determined to make a difference in a more significant way. It is this ardour that has driven me towards this research. Perhaps my own worst critic, I have never been content with what I have achieved, and I have always endeavoured to do better. I focus on acquiring new skills and seek out the best teachers I can find in my efforts to be the best Scientist Practitioner I can be. It is perhaps not surprising that this is the approach I take to this research.

1.2.4 Current role and observations

I am the Chief Clinical Officer of an online Cognitive Behavioural Therapy (CBT) provider called Ieso Digital Health (see www.iesohealth.com). The service provides CBT, via synchronous written (typed) communication to National Health Service (NHS) patients across England, within the Improving Access to Psychological Therapy programme (IAPT). This method of delivering CBT is discussed in greater detail in Chapter 3 of this thesis. Ieso Digital Health is private company that is venture capital funded. The company are contracted by the National Health Service (NHS) to deliver CBT to patients, across England, as part of the IAPT programme. Ieso Digital Health receives approximately 2,500 referrals each month and provides between 400 and 500 therapy appointments every day, 365 days a year[1]. These patients are treated by a team of 660 cognitive behavioural therapists who are managed and supervised by fourteen senior clinicians, clinical supervisors and clinical tutors. I have overall responsibility for the quality of care that is delivered to patients, the training and supervision provided to therapists and for research and development activities as well as budgetary and line management responsibilities. I have worked with Ieso Digital Health for seven years and started when the company first began treating patients in the NHS. I was the first clinician to join the company and the clinical policies, procedures and methodology have been driven by me. Having worked in many mental health care settings; delivering CBT, managing services, providing clinical supervision and teaching the clinicians of the future at a number of universities, I consider myself to be an experienced clinician and fully conversant with the conceptual framework of CBT. This relates very much to the scientist practitioner stance discussed in section 1.2.2, above. The provision of CBT (particularly in the United Kingdom) involves using evidence-based treatment protocols and measuring patient progress using standardised validated outcome measures at every therapy appointment (Layard and Clark 2014). In addition, therapists are encouraged to be reflective practitioners (Bennett-Levy, 2006, Bennett-Levy and Thwaites, 2007) in order to

[1] Correct at the time of writing

learn from every patient that they treat and to use this learning to inform their continuing professional development. Cognitive behavioural therapists are required to identify skills or knowledge gaps and to seek guidance from more experienced clinicians in the form of clinical supervision, self-directed learning and further training (BABCP 2019, NHS England 2018). Furthermore, cognitive behavioural therapists are encouraged to regularly record therapy sessions and present these recording to their clinical supervisor for feedback and guidance (BABCP 2019). These professional principles and guidelines serve the function of maintaining and improving clinical outcomes and are central to evidence-based psychological therapy (Layard and Clark 2014). Despite this significant focus on measuring outcomes, reflective practice and presenting live examples of clinical work to a clinical supervisor, it is difficult to identify why some therapists get better outcomes than others. For the first time in my career, my role at Ieso Digital Health has enabled me to study why some therapists may be better than others. This is because the online written format of delivering CBT, developed by Ieso Digital Health, means that for the first time ever it is possible to have access to transcripts of every therapy session delivered by a therapist together with the outcome measures for that session. Whilst a huge privilege, access to this amount of data has provided me with the opportunity to study therapist behaviour with the aim of improving outcomes for patients.

1.2.5 Rationale for this research

Sections 1.2.2 and 1.2.3 of this chapter have presented the professional context, theoretical framework and personal motivation for this research. The decision to focus on understanding the relationship between therapist variables and clinical outcome was driven primarily by three factors:

1. My personal curiosity and drive to understand why some cognitive behavioural therapists are obtaining better clinical outcomes than others, despite the fact there is

an established national curriculum for training cognitive behavioural therapists and minimum training standards for all accredited CBT therapists (NHS England 2018).

2. The unique availability of the transcripts derived from the cognitive behavioural treatment of over 40,000[2] patients.
3. The identification of a significant gap in the literature relating to what is known about which therapist variables are related to clinical outcomes in England's Improving Access to Psychological Therapy programme (IAPT).

1.2.6 Defining the term 'therapist variables'

In this thesis I will use the term 'therapist variables' to mean the individual differences between therapists both in terms of the individual differences in their demographics (for example age, gender, years of experience) and the differences in the way therapists conduct CBT with their patients. These differences between therapists are one set of variables that can account for variance in clinical outcomes (Nissen-Lie, Monsen and Ronnestad, 2010). Therapist variables are independent of patient variables and service variables, both of which also impact on the variance in outcomes (Johns, Barkham, Kellett and Saxon, 2019). This focus of the research discussed in this thesis is therapist variables.

1.2.7 Defining the term 'therapist effects'

The term 'therapist effects' is used to describe how therapists influence patient outcomes (Saxon and Barkham, 2012). The study of therapist effects has, more recently, focussed on the statistical analysis of large data sets using multilevel modelling. The therapist effect is quantified as the proportion of which the variance in outcome is attributed to the therapist (Saxon and Barkham, 2012). This statistical method models the nested structure of the data, whereby patients are nested within therapists and therapist are nested in services. In this

[2] 40,000 patients had completed treatment at the time this research was conducted

thesis I refer to the body of 'therapist effects' literature in order to illustrate how other researchers have explored variance in outcomes particularly within the Improving Access to Psychological Therapy (IAPT) programme.

1.3 OVERVIEW OF THIS THESIS

The structure of this thesis provides a sequential progression through the research. The order and content of each chapter is summarised below.

Chapter 1: Introduction to the professional context and rationale for this research
This chapter presents a personal reflection of the researcher's rationale for conducting this research in the context of her professional background and current role. Chapter 1 introduces the researcher in her role as Chief Clinical Officer at the company Ieso Digital Health. The researcher describes her experience and her theoretical position that informed the work discussed in this thesis. The term Scientist Practitioner is defined both as it relates to the literature and to cognitive behaviour therapy. The researcher describes how the Scientist Practitioner stance has informed her work and how the overarching aims and objectives of this study came about. The problem that there is significant variance in clinical outcomes between IAPT services and between therapists is identified. The researcher outlines her interest in therapist variables and the term 'therapist variable' is defined. The concept of using a unique data set, consisting of the therapy transcripts of 40,000 patients, is introduced as a new method for studying the variance between therapists.
The chapter concludes by providing an overview of this thesis.

Chapter 2: The improving access to psychological therapy programme (IAPT):
Chapter 2 presents an overview of the United Kingdom's Improving Access to Psychological Therapy programme (IAPT). The chapter discusses how and why IAPT was developed and presents the key guiding principles and processes that are employed within the programme.

Chapter 2 presents how health policy has driven the adoption of digital methods of delivering CBT, focussing on the delivery of CBT using online written (typed) communication. This method is used widely in IAPT and is called Internet Enabled CBT (IECBT). The delivery of CBT via IECBT is central to this research in that it provides Clinical Supervisors and senior clinicians with access to therapy transcripts. The transcripts are used to assess therapist competence with the aim of enabling therapists to reflect on and learn from their practice. Chapter 2 concludes by presenting the outcome data that is reported to NHS England from IAPT. This chapter identifies that there is a significant variance in outcomes between services and between therapists. The problem of variance in clinical outcomes is used to inform the questions that are used in the literature review.

Chapter 3: Literature Review:

Chapter 3 commences with a preliminary scoping review of the wider literature relating to the therapist variables that might be associated with clinical outcomes. This includes the historical context of research in this area and the literature from other psychological therapy models, including English speaking countries other than the United Kingdom. This is followed by a systematic review of the literature relating to peer reviewed papers which examine the relationship between therapist variables and clinical outcome in the United Kingdom's IAPT programme. The systematic review identifies a small number of papers which primarily focus on therapist effects in IAPT using Multi-Level Modelling. It is clearly established that between 3 and 8% of variance in outcome is due to therapist effects in IAPT. However, a significant limitation of all the papers reviewed is that most of the studies were unable access to recordings of therapy sessions or therapy transcripts and, of those that did, access was significantly limited. This chapter argues that whilst variance in outcome, due to therapist effects, is being established, very little is known about how therapists vary. It is identified that, up until now, it has not been possible to understand what therapists are doing with their patients and how that might relate to outcome. This chapter

identifies a significant gap in knowledge relating to what is known about therapist variables and their relationship with outcome.

Chapter 4: Methodology:

This chapter establishes the research design and methods used for this research and provides a justification for this approach in the context of the findings from the literature review and the ontological and epistemological positions that are presented in this thesis. Chapter 4 defines how the data for this naturalistic observational study was collected as part of normal routine care in the delivery of CBT by the service Ieso Digital Health. The storage and protection of the data, including ethics and data security are defined. The chapter describes how therapists who had delivered CBT to more than 10 patients were invited to allow the researcher to use their data for the purposes of this research. This is followed by a detailed description of the therapist data including therapist demographics and an analysis of each therapist's ability to deliver CBT with fidelity to the model and adherence to the evidence base. The terms 'F score' and 'A score' are introduced relating to fidelity the CBT model and adherence to the evidence base respectively. Detailed descriptions of the processes that were involved in assessing therapists' F scores and A scores including the process of inter-rater reliability training for each of the senior clinicians who rated the therapists' work. Chapter four concludes by outlining the statistical analyses plan for the data collected.

Chapter 5: Findings:

Chapter 5 presents the data produced from the rating of therapists' transcripts, therapist demographics and the associated outcome data. The chapter commences with a description of the dependent and independent variables and establishes normal distribution. This is followed by simple correlations and linear regression used to understand the relationship between each of the independent variables and the dependent variable; clinical outcome. It is established that there is no significant relationship between therapist

demographics and clinical outcome. This confirms the findings from other studies. However, therapist competence (as rated by the CTS-R) and therapist adherence are significantly related to clinical outcome at the <0.05 level. To further test the strength of this relationship, in the context of all the other variables, a hierarchical linear regression model was used resulting in similar findings in that competence and adherence remain statistically significant. A final analysis included patient variables into a Log-Linear Analysis model in order to understand whether therapist competence and adherence remain significant when patient variables are added to the model. The findings from this final analysis confirm the previous results in that both competence and adherence remain significant, but that adherence only had a relationship with clinical outcome through competence. This would suggest that therapist competence (as measured by the CTS-R) might be the vehicle from which the evidence-based protocol is delivered. That is to say if a therapist is unable to deliver CBT with fidelity to the model then it follows that they will be less likely to be able to adhere to a protocol and, without competence, the outcome is likely to be poor.

Chapter 6: Discussion:

This chapter presents the main findings produced by the research and discusses why the results and findings may have occurred and how these relate to what is already known. The chapter defines whether each research question was affirmed and discusses how the findings relate to the existing literature. This chapter argues that the findings from this research confirm the cited hypotheses in the literature that fidelity to the CBT model (therapist competence) and therapist adherence relate to clinical outcomes. It is also argued that the findings from this research confirm that therapist drift is a commonly occurring phenomena. Previously researchers had only been able to hypothesise that therapists may drift away from delivering CBT with their patients. It is argued that this finding has an impact on qualified IAPT therapists and for the training and assessment of trainee therapists. This chapter includes a discussion about the limitations of this research including a broader discussion of whether the findings from this research can be generalised to traditional face-

to-face CBT. Chapter 6 also explores that whilst competence and adherence a significantly related to outcome at the < 0.05 level, further research is required to understand what other factors may be related to outcome and this may require new methods of assessing competence or redefining what competence actually means. Whilst it is acknowledged that further research is required, this chapter concludes by asserting that the findings from this research have significantly contributed to the knowledge in the field. It is argued that the research presented in this thesis uses a new method of conducting psychological therapy research. The ability to study therapy transcripts in volume provides a new and innovative way to study what therapists do with their patients. Chapter 6 discusses how the findings from this research may be used by cognitive behavioural therapists, clinical supervisors, clinical services, higher education settings and policy makers. This chapter argues that, whilst the findings from this research suggest that competence and adherence are important factors for both qualified clinicians and those in training, the greatest impact, in the first instance, might come from implementing changes in the researcher's own service. It is argued that, in the last seven decades, psychological therapy research has predominantly focussed on demonstrating that one psychological therapy is superior to another. Despite the large numbers of published randomised controlled trials very little is known about the active ingredients of CBT and how it works. This chapter argues that the research reported in this thesis represents a new way to conduct psychological therapy research in the form of digitally delivered therapy data collected and analysed via the Ieso Digital Health platform. This data may be studied by clinicians, clinical scientists and Artificial Intelligence Scientists in order to understand the mediators and moderators of CBT as well as understanding what works for whom. The chapter concludes with an outline and rationale for the research and development projects and consequential service developments that have been implemented as a result of this research described in this thesis.

This chapter concludes that this research has extended the work of others and confirmed the previously surmised supposition that fidelity to the CBT model and adherence to an evidenced protocol are significantly related to outcome in IAPT.

Chapter 7: A personal reflection on the implications of this research
This final chapter presents the researcher's personal reflections on undertaking this research. This includes a discussion about real-world research and the tensions related to being an 'insider researcher.' Chapter 7 includes a reflection on how the findings have impacted on the researcher's own professional practice with a particular focus on the revelation that despite the fact that the researcher has significant experience in the field of CBT, as a clinician, teacher and senior leader, it has become clear that there is so much more to learn. The chapter concludes by defining the research questions that have developed since the research described in this thesis has concluded. These include investigating whether it is possible to teach a consistently poorly performing therapist to become a better performing therapist, investigating whether automated versions of the CTS-R might provide more effective ways of measuring therapist competence and exploring the effectiveness of clinical decision support tools.

1.4 CONCLUSION

The aim of this chapter was to provide a rationale for this research in the context of my professional role as Chief Clinical Officer of the online Cognitive Behavioural Therapy service, Ieso Digital Health, which is delivering CBT to patients across England as part of IAPT programme. The chapter has established the focus for this research and has presented some of the personal and professional beliefs and values have led to an interest in understanding what CBT therapists are doing with their patients and how this may account for the variance in outcomes between therapists. The following chapter presents an overview of the IAPT programme.

CHAPTER TWO: THE IMPROVING ACCESS TO PSYCHOLOGICAL THERAPIES (IAPT) PROGRAMME

Chapter One of this thesis presented the personal professional context of this research including an overview of this thesis. This chapter will consider England's Improving Access to Psychological Therapy (IAPT) programme which is the context for this research. A brief history of the development of IAPT will be presented, followed by an overview of the guiding principles and processes adopted by IAPT, including how health policy is driving the use of digital and online methods of delivering therapy in order to increase access to IAPT services. The use of digital methods of delivery will be explained with a particular focus on the Internet Enabled CBT (IECBT) method which is the research setting for this research. This chapter will highlight both the success of IAPT in increasing access to therapy and the emerging problems of training a large workforce to deliver CBT according to a set of evidence-based guidelines. It will be argued that one of the major problems of IAPT is the variance in outcomes between services and between therapists.

2.1 INTRODUCTION

Common mental health disorders, such as anxiety and depression, account for 38% of the disease burden in Western Europe (World Health Organisation, 2008). That is more than cardiovascular disease, cancer and diabetes combined. At least one in four people experience a mental health problem in any year (Layard Report 2012).The last Adult Psychiatric Morbidity Survey published in 2016 estimated that 5.9% of the UK population suffered from Generalised Anxiety Disorder and 3.3% suffered from depression (NHS Digital 2018a) It is estimated that common mental health disorders, such as anxiety and depression, cost the United Kingdom (UK) approximately £105 billion a year (No Health Without Mental Health 2011). The World Health Organisation (2001) estimated that, by the year 2020, depression will be the second most common illness requiring treatment and

intervention. Disorders such as depression and anxiety cause significant suffering and distress and are often disabling, affecting a person's ability to undertake activities of daily living or work (Layard, 2017). The consequential financial burden to UK government results from absenteeism from employment, reduced productivity in the workplace, frequent attendance in primary and secondary healthcare settings and reliance on disability benefits. The Organisation for Economic Co-operation and Development (OECD, 2012) report that over one third of all disability benefits in the UK are paid to claimants whose primary disorder is anxiety or depression. Layard (2017), an economist, estimates that the total cost to the UK is to reduce the country's gross national product (GDP) by 7%. This is more than the UK spent on education in 2017 which equated to 4.4% of the GDP (OECD 2018). Layard (2017) argues that, despite the overwhelming evidence of the prevalence of anxiety and depression and the cost to the country, the UK government spends only 1% of GDP on the provision of health care for all mental health disorders. He argues that providing effective treatments for people who present with anxiety and depression would increase the GDP by 4%. This argument was the motivating factor that led to Layard and Clark (2014) lobbying the UK government to increase access to effective treatments for patients with common mental health disorders. IAPT was clearly driven by an economic argument in the backdrop of a Brown/Blair, New Labour government. Layard's argument that mental health had traditionally been underserved, within an NHS that valued parity and equality of access, was arguably well-timed (Campling, 2019).

2.2 THE DEVELOPMENT OF THE IAPT PROGRAMME

Clark and Layard (2014) describe how they lobbied the UK government with their compelling argument that increasing the expenditure on treatments for common mental health disorders would not only be cost-neutral but had the potential to make the UK more profitable overall. Clark, a Clinical Psychologist and Layard, an economist, argued that a new type of nationwide provision for the effective treatment of anxiety and depression was required. Clark believed that the new service should draw on the effectiveness of Cognitive

Behavioural Therapy (CBT). Using evidence drawn from the Cochrane Collaboration (Clark, 2018) and the National Institute of Health and Care Excellence (NICE, 2004a, 2004b, 2005a, 2005b, 2006, 2009a, 2009b, 2011) it was argued that CBT held the widest evidence base for the treatment of anxiety and depression. This focus on evidenced-based interventions such as CBT has arguably led to a disinvestment in other forms of psychotherapy, such a counselling and psychodynamic psychotherapy (Guy, Loewenthal, Thomas and Stephenson, 2011). This is the antithesis of the United Kingdom government's promise (Department for Health and Social Care, 2011) that people should have a choice in relation to what mental health treatment they have. Furthermore, there is somewhat of a paradigm war between evidenced-based psychological therapies, and counselling and psychodynamic psychotherapy whereby the former favour 'gold standard' randomised controlled trials and outcomes-based frameworks and the latter prefer qualitative research and are less likely to use standardised validated instruments (Mollon, 2010). Furthermore, the National Institute of Health and Social Care Excellence (NICE) favour a biomedical approach to research supporting outcomes-based frameworks, randomised trials and experimental designs (NICE, 2017) and, therefore, psychological therapy models that do not share these paradigmatic beliefs are unlikely to become NICE approved (Guy, Loewenthal, Thomas and Stephenson, 2011, Mollon, 2010). Despite the counter arguments that alternative paradigms are useful in understanding what works for whom, the National Institute of Health and Care Excellence (NICE) reviews the evidence for effectiveness, for common mental health disorders such as anxiety and depression. Their reviews lead to the recommendations of particular disorder-specific treatment protocols using a cognitive behavioural framework, as well as some recommendations for other types of therapy, including counselling, but for depression only. The recommendations were based on the recovery rates reported in large scale randomised controlled trials (Clark ,2018). Layard, Clark, Knapp and Mayraz (2007 p.7) argued that the "new programme would pay for itself within five years" resulting from recovery rates of 50%, reduced prescribing, reduced absenteeism and increased productivity. The dominance of evidenced-based psychological interventions, particularly CBT, might be questionable

(Gaudiano, 2008). Although CBT is arguably the most evidence-based psychological intervention, it does not follow that it is the most effective or that it is equally effective for every patient (McPherson, Evans, Richardson, 2009). Clinicians from other psychotherapeutic modalities, such as counselling or psychodynamic psychotherapy, argue that the validated outcome measures used to quantify clinical outcomes are not fit for purpose and that there are more appropriate ways of measuring success such as asking patients whether they have achieved what they wanted to achieve (Friedi and Stearn, 2015). Furthermore, it might be argued that the evidence base for CBT has been over-represented (Wampold, Fluckiger, Del re, Yulish, Frost, Pace, et al., (2017). Despite the impotent counter arguments against CBT (Campling, 2019), Clark and Layard were successful in lobbying the government, leading to the development of the Improving Access to Psychological Therapy (IAPT) programme. The programme was designed to be a large-scale psychological therapy service, providing evidence-based interventions for people with anxiety and depression (Clark, 2018). Services were rolled out across England between 2008 and 2011 and were overseen by the Department of Health (Clark, 2012). From 2011 each area of England was required to have developed an IAPT service. Consequently, every General Practitioner in England could access evidenced based psychological treatments, such as cognitive behavioural therapy, for their patients. Since 2013, responsibility for IAPT services was devolved to NHS England with a focus on parity of esteem with physical health care provision (Health and Social Care Information Centre, 2013). This new focus was heavily influenced by the Health and Social Care Act (2012) which put legislation in place to ensure that mental health service provision was on a par with physical health care provision. However, it is important to note that, at its inception, IAPT aimed to treat just 15% of those people who presented with an anxiety disorder or depression (Clark, 2011) and that was only recently increased to a target of treating 25% of prevalence (Clark, 2018). Therefore, despite the change in policy relating to parity of esteem IAPT has failed to achieve this (The Guardian, 2014). It is questionable whether the general public would tolerate a health service that treated only 25% of people who have coronary heart disease, or diabetes and

yet IAPT celebrates (Clark, 2018) the fact that it is treating as many people as it is. Nevertheless, IAPT was designed to ensure, regardless of location, equal access to evidenced based interventions within a primary health care setting. Prior to IAPT access to evidence-based interventions was scant and availability varied across the country (Clark, 2012). Undoubtedly IAPT was commissioned by the government in order to make economic savings. Norman Lamb (at that time Minister of State for Care) argued that 45,000 people had come off state benefits because of the treatment they received via IAPT (NHS England, 2014). This focus on cost savings, to the United Kingdom's Treasury, may be a double-edged sword in that, whilst depression and anxiety can interfere with a person's ability to work or seek work (Mental Health Foundation, 2012), some of the socio-economic factors associated with anxiety and depression are less amenable to psychological therapy. The Department for Work and Pensions (DWP), in association with IAPT, sought to further the pursuance of reducing the economic burden of unemployment due to common mental health disorders such as anxiety and depression (van Stolk, Hofman, Hafner and Janta, 2014). The suggestion that claimants with a diagnosis of anxiety or depression should be offered CBT is also a double-edged sword. On the one hand, who could argue that treatments should not be offered to people who need them but ,on the other hand, defining employment as a health outcome might be more problematic (Elliott, 2018, Friedi and Stearn, 2015). If employment is a health care outcome, then it might be argued that therapists feel disempowered to enable people to achieve that aim. Additionally, if people are aware that the returning to work or obtaining a job is a desired outcome of therapy then they may feel disinclined to attend therapy (Campling, 2019; The Guardian, 2015). Whilst the overarching aims and objectives of IAPT are undoubtedly positive, it is important to note the tensions that arise when a service is established on the basis of cost savings. The development of IAPT reflected the clinical policies already established by the National Institute for Clinical Excellence (NICE) and a series of guidelines, advocating the use of disorder specific, evidenced based, psychological therapy protocols (NICE 2004a, 2004b, 2005a, 2005b, 2006, 2009a, 2009b, 2011) were published. It is important to note that, before IAPT, these

protocols were not routinely used in NHS services and, other than the randomised controlled trials, there was very little real-world evidence that they would be effective, at scale, in clinical practice (Gyani, Pumphrey, Parker, Shafran and Rose, 2012). Despite this, IAPT defined that therapists should be trained to use them and that this would lead to 50% of patients recovering (Layard and Clark, 2015). Under Clark's leadership IAPT set out reduce the financial burden of common mental illness by enabling people to return to work, reduce the prescribing of psychotropic medication and improving the nations wellbeing (Clark, 2012). The IAPT programme has three guiding principles (NHS England, 2017):

1. **The provision of a stepped care model**. Evidence-based psychological therapies are delivered by qualified and accredited clinicians using the most effective, but least burdensome (to the patient), treatment protocol.
2. **The use of routine clinical outcome monitoring**. Validated and reliable outcome measures are used to monitor each individual patient's progress. This data is collected at every appointment and reported, anonymously, to NHS England.
3. **The mandatory use of outcomes focussed clinical supervision**. Clinical supervision is provided, on a weekly basis, by more experienced clinicians and supports therapists to continuously enhance and improve their clinical practice.

2.2.1 The Stepped Care Model

The stepped care model (see figure 2.1) was introduced by NICE in 2004 as a recommended policy for the treatment of anxiety and depression (Clark, 2011). The model shows how patients are stratified into one of five steps according to the severity of their symptoms and their clinical needs. As such, the stepped care model was firmly adopted and adhered to in the development of the IAPT programme in 2008. All IAPT services use the stepped care model to treat patients at steps two and three. The stepped care model places a focus on ensuring that patients are offered an evidenced based intervention that is not

more intensive than required (Seekles, Straten, Beekman, Marwijk and Cuijpers, 2011). Following this model, many IAPT services allocate patients to a low intensity (LI) intervention, at step 2, first. Low intensity interventions are provided by Psychological Wellbeing Practitioners and patients are generally not offered more than 6 sessions. Low intensity interventions place an emphasis on providing patients with psycho-education about presenting symptoms followed by advice and guidance on how to manage symptoms more effectively (Seekles et al.,2011). These interventions may be provided face-to-face, over the phone, in groups, via email or through web-based self-help programmes. Patients at step two are routinely monitored and if they fail to respond to treatment they are 'stepped up' to a more intensive treatment at step three (Clark, 2018). Step three interventions are delivered by qualified high intensity therapists using an evidence-based intervention such as CBT or counselling for depression (NHS England 2016).

All IAPT services use the stepped care model, but it is often open to interpretation (Gellatly, 2011). Some services use the stepped care model sequentially so that all patients start at step 2, others place patients on a step according to the severity of their condition. Whilst the stepped care model has been demonstrated to be cost effective (Gellatly, 2011), and cost savings are implicit with the IAPT model, this might not always be to the patients benefit (Delgadillo, Gellatly and Stephenson-Bellwood, 2013).

The stepped care model relies on service policy and practice and individual therapist decision making. Either may be flawed, with services placing all patients at step two (the cheapest option) and individual therapists failing to step a patient up when it is clear that they are not benefiting from treatment at step 2 (Delgadillo et al., 2013). In either case patients fail to access an intervention that is most likely to help them, and this is the antithesis of the aims of the stepped care model (Clark, 2011) which ensures services provide the least intrusive, most effective intervention first, do not use single criteria (such as symptom severity) to determine movement between steps, and monitors progress and

outcomes to ensure the most effective interventions are delivered and the person moves to a higher step if needed.

Figure 2.1 The Stepped Care Model (NICE, 2011) showing how patients are classified according to severity/complexity of their presentation

The stepped care model

The recommendations in this guideline are presented within a stepped care framework that aims to match the needs of people with depression to the most appropriate services, depending on the characteristics of their illness and their personal and social circumstances. Each step represents increased complexity of intervention, with higher steps assuming interventions in previous steps.

Step 1: Recognition in primary care and general hospital settings

Step 2: Treatment of mild depression in primary care

Step 3: Treatment of moderate to severe depression in primary care

Step 4: Treatment of depression by mental health specialists

Step 5: Inpatient treatment for depression

	Who is responsible for care?	What is the focus?	What do they do?
Step 5:	Inpatient care, crisis teams	Risk to life, severe self-neglect	Medication, combined treatments, ECT
Step 4:	Mental health specialists, including crisis teams	Treatment-resistant, recurrent, atypical and psychotic depression, and those at significant risk	Medication, complex psychological interventions, combined treatments
Step 3:	Primary care team, primary care mental health worker	Moderate or severe depression	Medication, psychological interventions, social support
Step 2:	Primary care team, primary care mental health worker	Mild depression	Watchful waiting, guided self-help, computerised CBT, exercise, brief psychological interventions
Step 1:	GP, practice nurse	Recognition	Assessment

2.2.2 The Use of Routine Clinical Outcome Monitoring

The routine use of validated outcome measures at every therapy appointment is integral to the IAPT model (NHS Digital, 2016). Measures are used to confirm or disconfirm a patient's diagnosis and are also used as a measure of clinical improvement. Pre and post-intervention outcome measures of patients are used to calculate recovery or clinically significant improvement (the terms recovery and reliable improvement are discussed in more detail in chapter four). The Patient Health Questionnaire (PHQ-9) and Generalised Anxiety Disorder Questionnaire (GAD-7) form part of the mandatory minimum data set and are routinely used at every appointment (NHS Digital, 2016). In addition, a therapist may select further, anxiety disorder specific measures (ADSM), if the PHQ-9 and GAD-7 fail to be sensitive to the patient's presenting problem.

<u>GAD-7</u>

The GAD-7 (Spitzer et al., 2006) is a seven-item measure for anxiety using a 4-point Likert scale (0-3, where 0 indicates the absence of a symptom and 3 indicates greater severity). A cut off point of ≥ eight indicates greatest sensitivity and a clinical case in a primary care population. A cut of ≥ 15 indicates severe symptoms.

<u>PHQ-9</u>

The PHQ-9 (Kroenke et al., 2001) is a nine-item measure for depression using a 4-point Likert scale (0-3, where 0 indicates the absence of a symptom and 3 indicates greater severity). A cut off point of ≥ ten is used to indicate a diagnosis of depression in a primary care population.

The IAPT outcome-based framework is discussed in further detail in chapter 4 of this thesis.

2.2.3 The mandatory use of outcomes-based supervision

It is widely accepted that the provision of regular clinical supervision is necessary to support the continuing professional development of cognitive behavioural therapists (Milne, 2008, Lomax, Andrews, Burruss and Moorey, 2005 and Padesky, 1996). Turpin and Wheeler (2011 p.6) describe clinical supervision, in the context of IAPT, as:

"…*a formal relationship in which there is a contractual agreement that the therapist will present their work with clients in an open and honest way that enables the supervisor to have insight into the way in which the work has been conducted*…"
(Turpin and Wheeler, 2011 p.6)

Turpin and Wheeler state that the primary functions of clinical supervision are to ensure that the therapist is providing safe and effective clinical interventions, optimise clinical outcomes, build clinical skills and facilitate reflection on clinical practice. In light of the evidence that clinical supervision is likely to lead to enhanced practice, and therefore better clinical outcomes (Roth and Pilling 2008, Milne and James, 2000, Worthern and Lambert, 2007 and Miller, Duncan, Brown, Sorrel and Chalk 2007), IAPT provide detailed guidance on how supervision should be provided in every IAPT service in England. The guidance stipulates that each therapist will receive weekly supervision for at least one hour and that all of the therapist's patients should be discussed at least monthly. In addition, the IAPT guidance suggests that clinical supervisors should be suitably qualified and experienced and must have attended specific training in the provision of CBT clinical supervision (Turpin and Wheeler 2011). The IAPT programme stresses that the provision of high-quality supervision in accordance with its guidelines is essential to the overall success of the IAPT programme.

2.3 BUILDING THE IAPT WORKFORCE

The first IAPT services were developed in 2006 with two initial pilot sites in Doncaster and Newham (Richards and Suckling, 2009). Both pilot sites collected clinical outcome data using the PHQ-9 and GAD-7 at every appointment and 9 months after the last appointment (Clark, 2012). In addition, the pilot sites collected data that related to the patients' employment status. Layard's (2007) economic argument for the development of the IAPT programme was that the service would enable patients to return to work and that 50% of the patients treated would recover. The outcomes from both sites exceeded expectation in that there was 5% improvement in employment (those who were not working at the start of treatment but were working at the end of treatment) and the overall recovery rate was 52% (Clark, 2012). Following the perceived success of the IAPT pilots in Doncaster and Newham the UK government supported a phased roll out across England with a ring-fenced budget of £309 million (Steen, 2019). It was widely recognised (Department of Health, 2011) that there was a lack of a suitably qualified and experienced workforce to deal with such large numbers of patients. Therefore, the government set aside a proportion of the ring-fenced money to establish post-graduate clinical treatment training programmes at universities in the UK. The university programmes, commissioned by the NHS, focus on a Cognitive Behavioural model and provide specialised clinical training at two levels (Clark, 2012).

2.3.1 Psychological Wellbeing Practitioners (PWP)

PWPs are trained to provide low intensity interventions to patients at step 2 in the stepped care model (Clark, 2011). Low intensity interventions place a focus on enabling patients to work through a self-help treatment programme. The treatment may be based on bibliotherapy (written materials given to the patient), face-to-face interventions, telephone interventions, computerised self-help materials (cCBT) or online, guided self-help where the PWP supports the patient to use web-based self-help materials. Both cCBT and online guided self-help are discussed later in this chapter.

2.3.2 High Intensity CBT Therapists

High Intensity cognitive behavioural therapists are trained to provide one-to-one CBT using disorder specific treatment protocols (Layard and Clark, 2014). High Intensity Interventions are typically used where a patient has not improved after accessing a step 2 (low intensity) intervention or has a more complex or sever presentation requiring more intense and in-depth treatment.

In the first three years of the IAPT roll out 2,160 High Intensity CBT therapists and 1,440 Psychological Wellbeing Practitioners were trained (Clark, 2012). In 2009, the IAPT programme was extended to provide other evidence-based psychological interventions including counselling for depression, Brief Dynamic Interpersonal Therapy for Depression (DIT), Interpersonal Psychotherapy for Depression (IPT) and Couple Therapy for Depression (Department of Health, 2009). Specific training was providing for High Intensity Therapists (HIT) to support this extended provision. At the time of writing, 10,500 psychological therapists have been trained by IAPT (Clark, 2018) with at least half of those being High Intensity CBT therapists. These therapists work in over 200 IAPT services across England, with a ratio of 40 therapists per population of 250,000 (Health and Social Care Information Centre, 2014). In 2018 2.01 million patients (NHS Digital, 2018) were seen in IAPT with the majority of patients (68%) receiving CBT at either step 2 or step 3 (Clark, 2018). High Intensity Cognitive Behavioural Therapists (HI CBTs) account for the majority of the IAPT workforce. IAPT mandate (Clark, 2012) that HI CBT therapists receive a minimum standard of training and are accredited by the British Association of Behavioural and Cognitive Psychotherapy (BABCP).

2.4 TRAINING HIGH INTENSITY CBT THERAPISTS

Training provision for High Intensity cognitive behavioural therapists is provided on

postgraduate training programmes at several universities across England. Each of the training programmes must be accredited by the British Association of Cognitive and Behavioural Therapy (BABCP). The BABCP stipulate that the minimum training standards to be considered for accreditation as CBT therapist are; attendance at a post graduate CBT training programme that consists of a minimum of 750 taught hours combining theoretical learning with clinical practice and, twice weekly, clinical supervision. The curriculum content is fully outlined in two key documents:

1. BABCP Core Curriculum Reference Document (Hool, 2010)
2. Improving Access to Psychological Therapies (IAPT) National Curriculum for High Intensity Cognitive Behavioural Therapy Courses (Department of Health, 2011).

These documents explicitly exemplify the curriculum content for all post graduate training programmes in the United Kingdom (UK). The curriculum mandates (Hool, 2010) that trainees are taught how to treat the following disorders using a range of disorder specific protocols:

1. CBT for Specific Phobia
2. CBT for Panic Disorder
3. CBT for Social Anxiety Disorder
4. CBT for Obsessive Compulsive Disorder
5. CBT for Post-Traumatic Stress Disorder
6. CBT for Generalised Anxiety Disorder
7. CBT for Health Anxiety
8. CBT for Depression
9. Cognitive Therapy for Depression
10. Behavioural Activation for Depression

Specific guidance is given in relation to which protocols are taught for each specific disorder (Clark, 2018). For example, for the treatment of depression the curriculum mandates that trainees are taught how to deliver Beck's (1979) Cognitive Therapy and also Martel, Addis and Jacobsen (2001) Behavioural Activation protocol. The selection of specific protocols is based on NICE guidance and Roth and Pilling's (2008) framework, which outlines the clinical competencies for each of the recommended protocols (the Roth and Pilling framework will be discussed in section 2.5 of this chapter). At the end of clinical training candidates must demonstrate a series of clinical competencies including the ability to:

- *"construct maintenance and developmental CBT conceptualisations for depression and anxiety disorders*
- *develop CBT specific treatment plans*
- *practice CBT with depression and anxiety orders systematically, creatively and with good clinical outcome*
- *deal with complex issues arising in CBT practice*
- *take personal responsibility for clinical decision making in straightforward and more complex situation*
- *demonstrate self-direction and originality in tacking and solving therapeutic problems*
- *practise as "scientist practitioners" advancing their knowledge and understanding and develop new skills to a high level*
- *demonstrate a systematic knowledge of the principles of CBT and the evidence base for the application of CBT techniques*
- *demonstrate a systematic knowledge of CBT for depression and anxiety disorders*
- *a critical understanding of the theoretical and research evidence for cognitive behaviour models and an ability to evaluate the evidence*
- *demonstrate an ability to sensitively adapt CBT, and ensure equitable access taking into account cultural and social differences and values"*

Department of Health (2011 p.2-3)

Full details of the IAPT curriculum can be found at: http://www.iapt.nhs.uk/silo/files/national-curriculum-for-high-intensity-cognitive-behavioural-therapy-courses.pdf

2.5 SUMMATIVE ASSESSMENT OF IAPT TRAINEES

IAPT trainees are summatively assessed throughout their training (NHS England, 2017). Summative assessment is defined as a test or examination (either written or practical) that usually comes at the end of a course (Boulet, 2008). The assessment defines whether or not a student has met the key learning outcomes or competencies for the course they are studying. The student is awarded a final mark and is said to have passed or failed according to the pass mark that has been established by the training institution. Section 2.4, above, outlined the core clinical competencies that High Intensity IAPT trainees must demonstrate by the time they complete their training. These competencies, or learning outcomes, are assessed via two methods; submission of three audio recording of live therapy sessions (NHS England, 2018) and submission of written work (case reports and essays).

2.5.1 Summative assessment via written work

Trainees on the High Intensity CBT training programme are required to submit written work in the form of two case reports and 4 academic essays (NHS England, 2017, Clark, 2018). Of the two case reports, one relates to the comprehensive assessment of a patient and the other relates to a description of an entire episode of care. The pass mark for each piece of written work is 50%.

2.5.2 Summative assessment via recordings of live therapy sessions

Trainees are required to submit three recordings of live therapy sessions during their training (Clark, 2018). Each individual session must be drawn from three separate patients whom

the trainee has treated. The trainee must have received supervision from an IAPT training programme Clinical Supervisor for each of the three patients. The three recordings are assessed using the revised versions of the Cognitive Therapy Scale (CTS-R). The CTS-R was devised as evidence-based view of best clinical practice in the delivery of CBT (Blackburn, James, Milne, Baker, Standart, Garland and Reichelt, 2001). This tool is widely used to assess competence of trainee therapists and accomplished therapists alike. It is also used to assess fidelity to the CBT model in research studies. The tool consists of twelve items, each of which are rated on a 0-6 scale where a score of 3 on each item is considered competent. This 0- 6 rating scale is based on the Dreyfus and Dreyfus (1986) scale used to assess clinical competence in clinical trainees. The scale has been adapted by Blackburn et al., (2001) so that it might adequately rate levels of competence in trainee and qualified CBT therapists. The scale (as shown in figure 2.2) is used to rate a clinician on each of the 12 items of the CTS-R. Figure 2.2 shows the 0-6 rating scale where a score of 0 would indicate incompetence (absence of the skill) and a score of 6 would suggest that the clinicians is an expert who has a consistently high performance even in the face of adversity. The twelve items of the CTS-R relate to specific cognitive behavioural processes or techniques which are deemed to be the basic CBT competencies (Roth and Pilling 2007). The twelve items of the CTS-R are shown in figure 2.3.

The CTS-R is used widely throughout CBT training, not only to assess a student's ability to demonstrate core competencies in CBT but also to teach students what is required in order to be considered competent as a cognitive behavioural therapist. Whilst training, trainee's receive regular practical skills workshops where they are taught how to deliver CBT with fidelity to the model as assessed by the CTS-R (www.exeter.ac.uk). It might be argued that the trainers who are delivering these workshops are 'teaching to the test'. This phenomenon has been described as a pedagogical process whereby students are regularly exposed to the summative assessment, so that it might be seen as signification test preparation

(Volante, 2004). This process has been criticised because it can hamper learning in that the student focuses only on the final summative assessment and not on learning core clinical skills (Volante, 2004). This issue will be discussed in more detail in Chapter Six.

Figure 2.2 Rating clinical expertise in CBT adapted from Dreyfus and Dreyfus (1986)

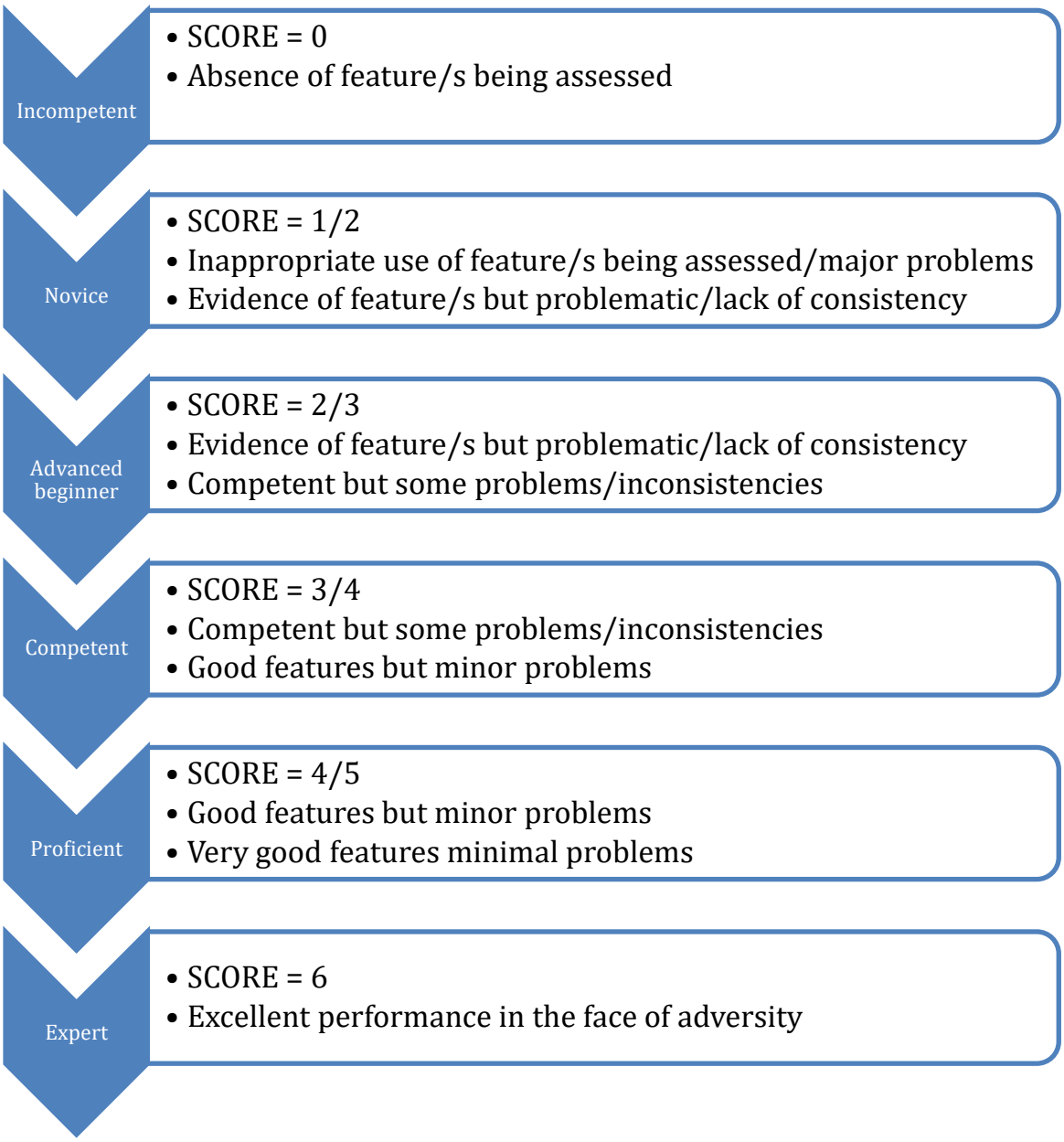


Figure 2.3 The 12 items of the CTS-R (Blackburn et al., 2001)

ITEM NUMBER	ITEM DESCRIPTOR	SCORE (0 -6)
1	Agenda setting and adherence	
2	Eliciting feedback	
3	Collaboration	
4	Pacing	
5	Interpersonal effectiveness	
6	Eliciting emotional expression	
7	Eliciting key cognitions	
8	Eliciting & planning behaviours	
9	Guided discovery	
10	Conceptual integration	
11	Application of change methods	
12	Homework setting	

2.5.3 Using the CTS-R to assess competence

The CTS-R is used as the summative assessment of therapist competence on IAPT training programmes (Clark, 2018). Blackburn et al., (2001) produced a manual that accompanies the CTS-R which provides guidance for each of the CTS-R items in terms of identifying examples where a therapist would be given a score 6, 5, 4, 3, 2 or 1. The manual is used as a tool to provide training to those new to marking and to guide groups of markers when undertaking inter-rater reliability training (Blackburn, James. Baker, Standart, Garland and Reichelt, 2001). Inter-rater reliability training is important as high inter-rater reliability can be hard to achieve (Blackburn, et al.,2001). It is necessary for teams of markers to become aware of their own subjective opinions and to receive training on how to identify quantifiable examples of competence on each item of the CTS-R (Blackburn et al., 2001).

High Intensity IAPT trainees submit three recording for summative assessment using the CTS-R. Trainees are required to achieve a score of ≥40% on the first summative CTS-R, followed by a score of ≥50% on the subsequent 2 CTS-Rs (www.exeter.ac.uk). Trainees submit recordings of therapy sessions to be assessed by the High Intensity IAPT training programme teaching team. Trainees self-select recordings on the basis that they perceive that the recording provides a good example of where they are demonstrating competency in relation to fidelity to the CBT model. Self-selection of recordings has been criticised as it is likely that the recordings are not wholly representative of the trainees' clinical work (Walfish, McAlister, O'Donnell and Lambert, 2012). This issue is discussed further in chapters six and seven of this thesis. Each submitted recording is rated, using the CTS-R, by two markers and then moderated by a third marker. Where the first two markers are unable to reach an agreement, the third marker makes a final decision. Where the trainee achieves the final pass mark for all three CTS-Rs they are deemed to be competent (Williams, Moorey and Cobb, 1991, Branson and Shafran, 2015, Liness, Lea, Nestler, Parker and Clark, 2016 and Clark, 2018). Whilst the CTS-R is used to assess competence on all IAPT training

programmes and is widely used in other settings too, its use as a tool to assess therapist competence could be debated. The original Cognitive Therapy Scale was developed as part of the National Institute of Mental Health Treatment of Depression Collaborative Research Programme (Vallis, Shaw and Dobson, 1986). Therefore, it was designed to assess therapist competence in the treatment of depression for a large research trial. It was not designed as a formative or summative assessment tool in Higher Education. The revised version of the CTS (the CTS-R) varies very little from the original version, although it was used by the authors of the tool to assess whether they demonstrated an improvement on the CTS-R as their training progressed (Blackburn et al., 2001). This may be the rationale for using it in Higher Education now, but its efficacy remains in doubt (Muse and McManus, 2013). Regardless of the arguments that suggest the CTS-R may not be the right instrument to assess therapist competence, it remains the instrument of choice pending the development and adoption of a more effective tool.

The full version of the CTS-R can be found as appendix item 6.

2.5.4 Using written work to assess adherence: Roth and Pilling (2008) competencies

While the assessment of trainee competence, using the CTS-R, is well established, the assessment of a trainee's ability to deliver CBT whilst adhering to a disorder specific protocol is less clear. Roth and Pilling (2008)[3] have provided very detailed clinical competencies for each of the disorder specific protocols taught on the IAPT training programme and these are used to inform the IAPT training curricula (Clark, 2018). However,

[3] Full details of the Roth and Pilling competency framework for each of the disorder specific protocols taught on the High Intensity CBT training programme can be found on the University College London website here: http://www.ucl.ac.uk/pals/research/cehp/research-groups/core/competence-frameworks

the summative assessment of adherence to any of the taught protocols is undertaken via the trainee's written work. One of the case reports that a trainee submits describes a whole episode of care with one patient and this is used, in part, to assess the trainee's theoretical understanding of that particular protocol. Additionally, trainees are required to submit essays on the evidence base for the treatment of anxiety disorders and depression and these too are used to assess the trainees' theoretical understanding of each protocol. At the current time there is no summative assessment of clinical practice in relation to a therapist's ability to apply their theoretical knowledge, in relation to each protocol, with patients. This thesis will argue that the premise that the efficacy of the evidence-based protocols would generalise to real-world settings and that it was possible to teach trainees to deliver the protocols on a 12-month training programme may be questionable. Chapters 6 and 7 of this thesis will explore the issue of the current methods used to assess trainee (and qualified) therapists' adherence to protocols and it will be argued that this may be one reason why there is a significant variance in outcomes between therapists.

2.6 EVIDENCED-BASED PSYCHOLOGICAL INTERVENTIONS

One of the aims of IAPT is to deliver evidence-based psychological interventions with a primary focus on clinical outcomes. Section 2.3.3, earlier in this chapter, outlined the training curricula for High Intensity CBT therapists. This curriculum focuses only on evidence-based psychological interventions (Layard and Clark, 2014). The movement towards evidence-based psychological therapies began in the late 1960's (McHugh and Barlow, 2012). This move away from anecdotal evidence of efficacy of treatment brought about the development of protocol-driven psychological interventions that can produce generalizable results across many populations (Kazdin, 2008). Since the 1960's, there has been a plethora of research relating to the efficacy of a range of psychological interventions. With a focus on a positivist approach and larger-scale experimental design studies it is widely reported that the most common psychological modalities such as cognitive behavioural therapy, systemic therapy

and psychodynamic therapy are safe and effective for large numbers of people (Nathan and Gorman, 2007, Coldwell and Bender 2007, Hofmann and Smits, 2008, Silverman, Pina and Viswesvaran 2008). McHugh and Barlow (2012) report that, by far, cognitive behavioural therapy (CBT) has the largest evidence base. They argue that the preference for large-scale outcome studies that are "tightly controlled" (McHugh and Barlow p. 4) has enabled CBT researchers to repeatedly demonstrate outcome benefits. However, Westbrook and Kirk (2005) earlier argued that whilst CBT had a strong evidence base that supported its application that there was dearth of evidence that demonstrated its application directly in clinical services. Westbrook and Kirk stated that most of the initial research into the efficacy of psychological interventions had been based in academic settings rather than clinical services. McHugh and Barlow (2012, p.7) describe this as the "research-practice gap." It might be argued that the gap between research and clinical practice can only be bridged by the development of more effective relationships between researchers and clinicians (Horsfall, Cleary & Hunt, 2011) and that evidence-based practice needs to be cultivated by, and for, clinicians. Evidence-based psychological interventions are not without their critics. One argument against evidence-based psychological interventions is that the evidence is not valid if it has been produced in an academic setting (Fisher and Happell, 2009). Whilst this is valid argument, it might be possible to bridge the gap between research and real-world clinical settings by focussing more of the role of scientist practitioners (Newnham and Page, 2010). Scientist practitioners may promote pragmatic trials in their own clinical settings. Pragmatic trials, unlike research trials in academic settings, provide valuable insights about whether research is generalisable in clinical settings (Holmqvist, Philips and Barkham, 2015) but despite Fisher and Happell's (2009) argument that pragmatic research is more valid, pragmatic research can also have its challenges. Recruiting therapists to treat patients in pragmatic studies can be problematic both because of the perceived additional burden on their workload but also because therapists may be fearful of the scrutiny of their work (Hatcher and Gillaspy, 2006). Further challenges present in relation to the interpretation of the findings from pragmatic studies both in as much as researchers may

lack objectivity and misrepresent their findings or the academic audience has a greater appetite for research conducted in academic settings (Homqvist, Philips and Barkham, 2015).

The Improvement Access to Psychological Therapy (IAPT) programme, whilst initially based on what Fisher and Happell (2009) would term 'academically sited research', supports the concept of research undertaken in the clinical setting. The postgraduate clinical training programmes commissioned by IAPT, in order to train its workforce, places an emphasis on trainees developing the skills of a scientist practitioner (Roth and Pilling, 2008). Pilecki and McKay (2013) describe the scientist practitioner model of clinical training as the provision of training in both research methods and clinical skills. They argue that this model equips clinicians with the skills necessary for academic enquiry and enhances their ability to become effective practitioners. The IAPT programme, led by David Clark, also places an emphasis on outcome research developed by economists, clinicians and clinical researchers. Clark, Layard, Smithies, Richards, Suckling and Wright (2009) report on the initial evaluation of the first IAPT sites and there have been annual publications of outcome data since then (Clark, 2012). Implicit within the IAPT model is the mandatory reporting of outcome data by all of the IAPT services. This data has been used by health economists to develop a statistical argument that supports the hypothesis that IAPT is able to meets its original aims and objectives of improving recovery, reducing prescribing and enabling people to return to work.

Whilst the published outcome data relating to IAPT services since 2009 has been largely positive (McHugh and Barlow, 2012) it is not without its critics. The We Need to Talk coalition present a counter argument in their 2010 research paper 'We Still Need To Talk.' The coalition, consisting of a number of high-profile organisations such as the Royal College of Psychiatrists, The Mental Health Foundation and The British Psychological Society, argue that IAPT is failing to provide an effective service. They state that IAPT has long waiting

times, that patients are not offered a choice and that many are not offered therapy at all. However, it cannot be argued that IAPT currently treats over half a million patients a year and many of these would have not received any treatment prior to IAPT (Clark, 2018).

2.7 IAPT IN 2019: CHALLENGES AND CHANGE

Despite increased funding and increased numbers of trained psychological therapists IAPT is failing to meet its targets (Strathdee, 2013). The majority of IAPT services are treating less than 11% of those that have anxiety or depression (We Need to Talk Coalition Report, 2010). The IAPT target is currently 15%. The Department of Health (2011) reports that health services in the United Kingdom (UK), are failing to respond to the growing need for evidence-based treatments such as Cognitive Behavioural Therapy (CBT). The 'We Need to Talk' coalition argue that one in ten people wait more than twelve months for treatment, more than half of patients must wait for more than three months for treatment, there is no choice of provision offered and there is a lack of parity of esteem with physical health provision. The coalition cite the government's commitment to parity of esteem in the Health Policy paper 'No Health Without Mental Health' (2012) and argue that the government is failing to address the problem. Davies (2014), in her Chief Medical Officer's Report, also acknowledges the deficit in provision within primary care mental health services. She cites emerging health policy that supports the use of technology to bridge the gap in the provision of evidenced based psychological interventions, such as cognitive behavioural therapy (NHS Mandate 2014, Hollis et al., 2014). The technological application of cognitive behavioural therapy (CBT) has been termed Computerised CBT (cCBT). This method of delivering cognitive behavioural interventions focuses mainly on self-help with no, or very minimal, therapist input. Computerised CBT enables patients to access self-help materials via a CD-ROM, computer software or on a website. Services of this kind include 'Beating the Blues' and 'Moodgym' both of which have been used with limited success in IAPT (Grist and

Cavannah, 2013). Therefore, it is unlikely that this type of intervention will provide the solution to bridge the gap in provision (Richards, Timulak and Hevey, 2012).

The history of technologically delivered cognitive behavioural interventions is a short one. Since the advent of evidence-based interventions, researchers have explored the efficacy of various methods of delivery. Initially, research focussed on face-to-face delivery and group intervention and, more latterly, telephone-, computer- and Internet-delivered interventions. The last decade has seen a growing body of evidence that demonstrates that some computer-based and Internet-based interventions can be effective (Hedman, Ljottson and Lindefors, 2012). Kanter et.al, argues that people can differ in their requirements from treatment. They suggest that not all patients will benefit from traditional face-to-face interventions and offering a range of options including self-help and online services may widen access and optimise the efficacy of treatment.

2.8 SUMMARY

This section has so far discussed the development of the IAPT programme in England. The achievements, challenges and counter arguments against the IAPT programme have been presented. An overview of the programme has been provided including the guiding principles of IAPT, the stepped care model, the focus on the provision of evidence-based psychological interventions, the outcome-based framework and the training curriculum for High Intensity CBT therapists. It has been argued that the assessment of trainees in relation to their ability to adhere to an evidence-based treatment protocol may be problematic in that it is currently impossible to assess what trainees (or qualified therapists) are doing with their patients. Sections 2.2, 2.6 and 2.7 have highlighted the growing voice of dissent against the IAPT programme. Those that have voiced a differing opinion are, most notably, from other (i.e. non CBT) theoretical and therapeutic traditions. The arguments presented include an opposition to the epistemic foundations of the IAPT programme, which have served to

exclude therapeutic modalities such as psychoanalysis, psychodynamic therapy and person-centred counselling. It has been argued that those that disagree with the way that the IAPT programme has been established have been impotent in their ability to effect change. Furthermore, clinicians and academics from traditions, other than CBT, have voiced concern that the IAPT programme is failing to tackle a growing increase in the prevalence in mental health conditions and that whilst IAPT reports that it is providing 'transparency' in reporting its data in the public domain, the data might not be as transparent as it is purported to be. It has been reported that this growing voice of dissent against the IAPT programme argues that Layard and Clark's promises to the government have been overstated, and the IAPT programme's focus on reducing government spend on benefits payments encourages coercive and anti-therapeutic practice amongst IAPT clinicians.

This section has concluded with a discussion regarding the difficulties that relate to widening access to evidence-based psychological interventions. Digital and online methods of delivering CBT are one way of widening access and are supported by health policy. Section 2.9, below, will discuss the online and digital methods currently used with the IAPT programme, focussing primarily on the online method Internet Enabled CBT (IECBT). This method of delivering CBT is the focus of the research described in this thesis.

2.9 DELIVERY OF CBT WITHIN IAPT: DIGITAL METHODS

Digital methods of delivering CBT are becoming more widely adopted and offer a number of key advantages to patients (Andersson, 2014, Andersson 2015, and Richards, Tumulak and Hevey, 2012). Patients often describe online interventions as more convenient and less anxiety provoking than face-to-face CBT (Andersson 2015). It is likely that online methods of delivering CBT enables patients, who might not otherwise be able to attend therapy appointments, to access treatment (Kessler et al., 2009). There is also evidence that suggests that patients often find it easier to discuss the nature of their problem when

working online because they feel less embarrassed or ashamed (Andersson 2014, 2015, Suller, 2015, Kessler et al., 2009). Additionally, online CBT has been found to be especially advantageous for specific disorders such as Social Anxiety Disorder, where patients find the social interaction in face-to-face CBT problematic (Stott et al., 2013). Furthermore, online methods widen access to evidence based psychological therapies, such as CBT because, unlike traditional face-to-face services, patients may access therapy on any day of the week, at any time and from where ever they choose (Andersson, 2015). These key advantages have led to changes in health policy whereby the NHS Long Term Plan (2019) place a significant focus on the adoption of digital methods in the NHS in general and more specifically to widen access to CBT both in IAPT and in secondary care settings (NHS England, 2019). Additionally, the latest IAPT manual gives specific guidance on how services should use online and digital method to widen access to patients (National Collaborating Centre for Mental Health, 2019). Online methods are firmly embedded in IAPT and digital methods of delivery have been established as recommended practice in mental health policies such as the NHS Long-term Plan (NHS England, 2019).

2.9.1 Brief historical context of online methods

The history of using technology to deliver cognitive behavioural therapy is a short one (Andersson, 2015). One of the earliest adopters of technology in the assessment and treatment of patients with common mental health disorders was Dr. John Griest, Professor of Psychiatry at the University of Wisconsin (Marks, Shaw and Parkin, 1998). Griest developed a computer delivered interview to predict suicidality amongst depressed patients (Griest et al., 1974). He argued that patients were far more likely to respond openly and honestly on a computer than if they were being assessed face to face by a doctor or health care professional. His studies in the 1970's demonstrated that a computer-based assessment was more effective in predicting the likelihood of a suicide attempt, than a risk assessment conducted by a psychiatrist (Kobak, Reynolds and Griest 1994). By the mid 1970's, Isaac

Marks, Professor of Psychiatry at Kings College, London became interested in Griest's work and began to investigate how technology might be used to optimize the delivery of evidence-based mental health interventions (Marks, Shaw and Parkin, 1998). Marks, in collaboration with Griest and his Wisconsin team, went on to develop computer-delivered, guided self-help materials, initially on CD rom and then later on web-based platforms. These materials were predominantly aimed at patients who presented with anxiety disorders such as specific phobia, post-traumatic stress disorder and obsessive-compulsive disorder (Marks, Shaw and Parkin, 1998). These early programmes included 'OBT', 'Fear Fighter' and 'OCBT'. Marks established a "cCBT Clinic" (Computerised Cognitive Behaviour Therapy) at the Maudsley Hospital in London, where local General Practitioners could refer their patients. Marks argued that not only was this new treatment 'just as good as face-to face treatment' but the service was able to operate without a waiting list because patients could access treatment on one of many computers at his clinic (Marks, 2013). Marks and Griest's work in this field has led to the development of a number of online and digital methods which are used to deliver CBT (Andersson, 2015). These methods fall into a series of categories on a continuum which ranges from self-help methods (no therapist involvement) at one end of the continuum and 100% therapist-delivered methods on the other end of the continuum. Figure 2.4 shows this continuum and illustrates where various methods sit on the continuum.

2.10 DIGITAL METHODS OF DELIVERING CBT IN IAPT

There is a growing vocabulary to describe various methods of digitally delivered CBT (Andersson, 2015). These include terms such as computer therapy, Internet interventions, media delivered CBT, e-therapy, telemedicine eCBT, IECBT and ICBT (Aboujaoude and Starcevic, 2014). This myriad of terms creates confusion when attempting to differentiate between the methods (Aboujaoude and Starcevic, 2014). At the time of writing there are predominantly four methods of delivering CBT in IAPT using digital methods. These are computerised CBT (cCBT), online guided self-help, virtual reality and the Ieso Method,

sometimes described as Internet Enabled CBT (IECBT). The following section will discuss each of these methods.

Figure 2.4 The continuum of online and digital products used to deliver CBT

2.10.1 Computerised Cognitive Behavioural Therapy (cCBT)

Computerised Cognitive Behavioural Therapy (cCBT) provides online self-help materials for patients (Andersson, 2015). These materials are usually on a web-based platform (website) or delivered via a smart phone app. Computerised CBT programmes have been developed for a wide range of mental health disorders including depression, anxiety disorders, chronic health conditions, sleep disorders and chronic pain (Andersson 2015). Other programmes provide psycho-education for general mental wellbeing such as mindfulness (see www.headspace.com). Marks, Kenwright, McDonough, Whitaker and Mataix-Cols, (2004), de Graaf et al., (2008), Berger et al., (2011) and Moritz et al., (2012), amongst others, have all argued that; cCBT is of benefit to patients (with moderate effect sizes), cCBT is highly cost effective as there is no therapist contact and cCBT is convenient and can be accessed by anyone with an Internet connection and a computer, tablet or (more latterly) a smart phone. The growing evidence base for cCBT encouraged the United Kingdom's National Institute for Health and Care Excellence (NICE) to endorse cCBT materials for the treatment of depression (National Institute for Health and Care Excellence 2013). One cCBT product

designed to treat depression known as 'MoodGYM' (Batterham, Neil, Bennett, Griffiths and Christensen, 2008 and Christensen et al., 2004) has been reported to have been used by over 400,000 depressed people across the world. Despite the evidence that cCBT can be effective for many common mental health disorders such as anxiety and depression there are a growing number of counter-arguments that relate to the reliability and validity of the findings from the cCBT studies (Aboujaoude and Starcevic 2014). One meta-analysis of the cCBT literature argued that the RCT's that had been conducted failed to demonstrate significant evidence of behavioural change and improvement in a patient's quality of life and that authors failed to rigorously report on patient drop out. (So et al., 2013). There is clear evidence that the drop-out rate from cCBT is extremely high with reports citing between 28% and 80% (Hilvert-Bruce at al., 2012), compared to a drop-out rate of approximately 20% in face-to-face therapy (Hans and Hiller, 2013). There is little doubt that the development of cCBT materials in the last 35 years has been driven by a desire to improve access to evidenced based interventions. This desire is situated within a climate of increasing prevalence of anxiety and depression, over-stretched health care budgets and a lack of qualified clinicians (Andersson 2015). It might be argued that the need to deliver more for less has led to some clinicians over-estimating the effectiveness of cCBT (So et al., 2013). On the other hand, it might be argued that even if cCBT is not as effective as was first thought that it is better than nothing for some people in parts of the world where there is little or no access to psychological treatments (Cuijpers et al., 2009). Regardless of this ongoing debate, the Improving Access to Psychological Treatment (IAPT) programme, continues to support the use of cCBT materials for patients who present with mild to moderate depression (NICE, 2013).

2.10.2 Online guided self-help (ICBT)

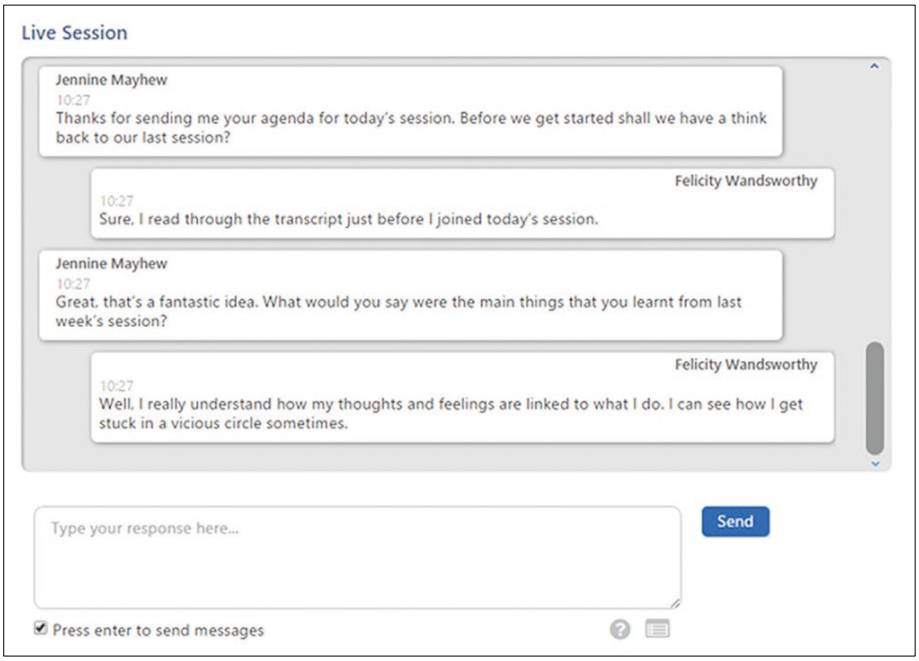
Online guided self-help (ICBT) is commonly used to treat patients at step 2 within IAPT (Gilbody et al., 2016). ICBT differs from cCBT in that patients are supported to use web-

based self-help materials (similar to those used in cCBT) with the help and guidance of a clinician (Andersson, 2015). The supporting clinician, in IAPT, is usually a Primary Care Wellbeing Practitioner (PWP). The support is provided by email (Kyrios et al., 2014), telephone (Zou et al., 2012), or occasionally face-to-face (Marks et al., 2004). Support can vary from between ten minutes and half an hour a week (Aboujaoude and Starcevic 2014). The aim of the support is to assist patients in working through the online material and therefore improve patient engagement. Whilst there are similarities between cCBT and ICBT, and often both methods use very similar online materials (e.g. MoodGYM), the evidence base for guided online CBT (ICBT) is larger (Andersson, 2014, Andersson 2015, Hedman, Andersson, Ljotsson, 2011, Cuijpers, Donker, van Straten and Andersson, 2010). There are over 50 RCTs relating to the efficacy of ICBT (Andersson, 2015). These studies relate to a wide range of disorders including; depression, obsessive-compulsive-disorder, social anxiety disorder, specific phobia, eating disorders, post-traumatic stress disorder, chronic pain and post-natal depression.

2.10.3 Internet Enabled Cognitive Behavioural Therapy (IECBT)

Internet Enabled CBT (IECBT) delivers CBT using synchronous written (typed) communication via a secure web-based platform (see www.iesohealth.com). Unlike computerised CBT and online guided self-help, IECBT is delivered by a qualified and accredited cognitive behavioural therapist, identical to face-to-face CBT. However, in IECBT the mechanism of communication is reading and writing, rather than speaking and listening. Unlike face-to-face CBT, the patient and therapist are unable to see each other (there is no video element to IECBT). The transcript of each therapy session, conducted in this way, is then held on the secure web-based platform for both therapist and patient to access at any time. An example of a CBT session conducted using this method can be seen in figure 2.5.

Figure 2.5 A (fictitious) example of an Internet Enabled CBT session



In addition to a weekly CBT appointment, therapist and patient can also communicate with each other in-between therapy appointments. This asynchronous communication can be used to amplify the effect of CBT by encouraging the patient to focus on out-of-session tasks and goals, and consolidating learning that has taken place during a therapy session. Contact with a therapist between appointments rarely occurs when CBT is delivered face-to-face. This in itself may have a positive effect on clinical outcomes.

IECBT has been demonstrated to be clinically effective and has broadly similar outcomes to those reported in face-to-face CBT (Kessler, Lewis, Kaur, Wiles, King, Weich, et al., 2009, Catarino, Bateup, Tablan, Innes, Freer, Richards et al., 2018). An early trial compared the delivery of IECBT with treatment as usual in 297 participants who were diagnosed with major depressive disorder. The participants were randomised to the intervention arm or the control arm (treatment as usual). Participants in the intervention arm had IECBT delivered by a CBT therapist. Treatment consisted of 10 sessions of CBT for depression, although the researchers do not state which depression protocol the therapists used. 113 participants that

were in the intervention group (that had completed treatment) and 97 in the control group were followed up four months after the intervention. In the intervention group 43 (38%) patients recovered from depression (Beck Depression Inventory score <10), versus 23 (24%) in the control group (odds ratio 2.39, 95% CI 1.23-4.67; p=0.011), and 46 (42%) versus 26 (26%) at 8 months (2.07, 1.11-3.87; p=0.023).Effect sizes at 4 months and 8 months were 0.81 and 0.70, respectively. Kessler et al., (2009 p.) concluded that: "CBT seems to be effective when delivered online in real time by a therapist, with benefits maintained over 8 months. This method of delivery could broaden access to CBT." The study demonstrated evidence of the concept of IECBT in that the recovery rate in the treatment arm of the study was equal to the published recovery rates of depressed patients in previous studies using face-to-face CBT (Kessler et al., 2009). In addition, IECBT has been used to deliver CBT within IAPT to over 40,000[1] patients by the company Ieso Digital Health. The outcomes for these patients are reported to NHS England and can be benchmarked against the reported outcomes for face-to-face CBT services. The average recovery rate for face-to-face services is reported as 52% and IECBT as 54% (Catarino et al., 2018, NHS Digital 2018). IECBT is being widely used within IAPT and, at the time of writing, approximately 2000 patients are referred for IECBT every month. Currently, no other provider is delivering CBT using IECBT in IAPT[2].

2.11 THE IECBT METHODOLOGY

The IECBT method was used to treat 14,000 IAPT patients in 2018. The overarching aim is to widen access to evidence-based psychological therapy (Catarino et al., 2018). However, it might be argued that the method is more than a therapy delivery platform, rather, it is a unique methodology with a specific focus on using data and technology in order to learn how to improve clinical outcomes. This methodology encompasses a series of processes which

[1] Correct at time of writing October, 2019
[2] Correct at the time of writing October, 2019

aim to support therapists to be the best they can be. This section will present the methodology that is in use at the time of writing.

2.11.1 Recruitment and selection of therapists

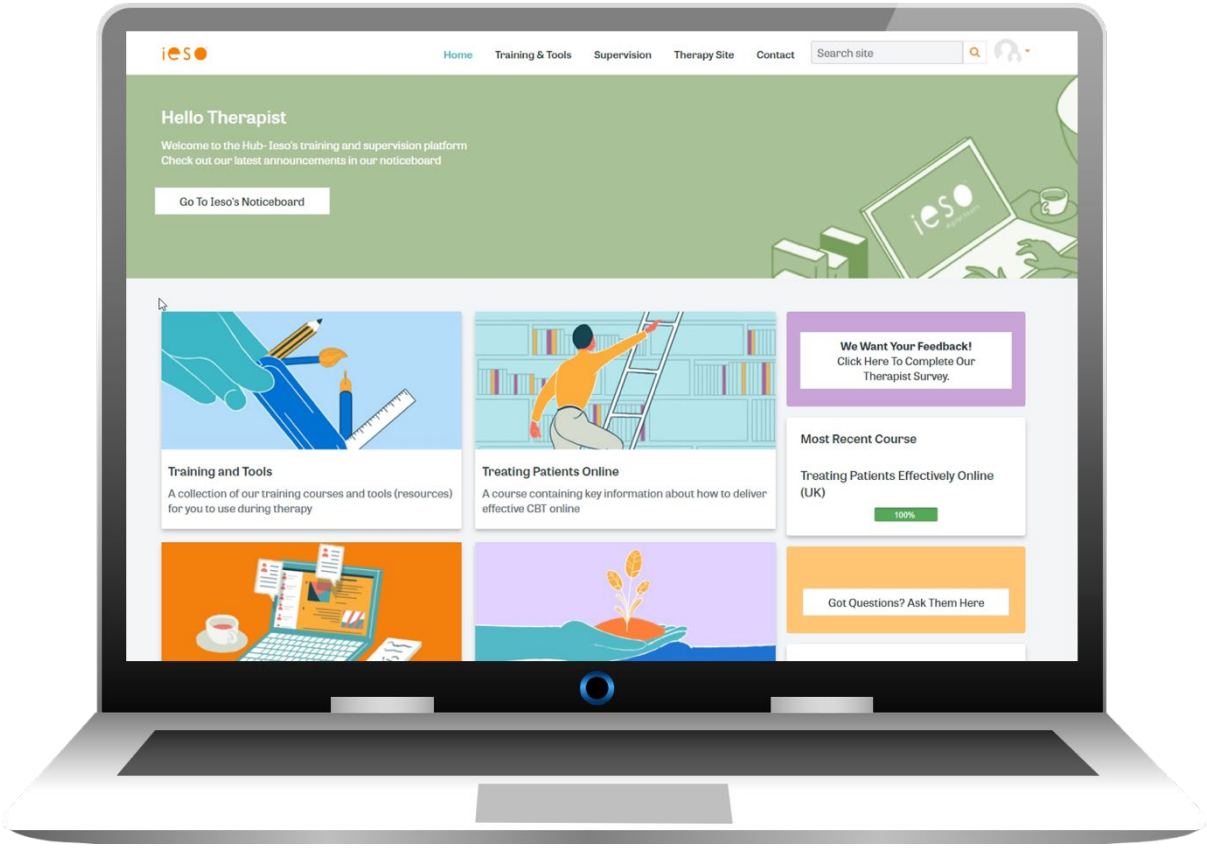
Like all IAPT services (Clark, 2014), IECBT (see www.iesohealth.com) recruits only British Association of Behavioural and Cognitive Psychotherapy (BABCP) accredited CBT therapists. Chapter two provides further detail relating to this accreditation process. Once recruited, therapists undertake a period of assessment whereby their work with patients on the IECBT platform is reviewed by a clinical supervisor. The supervisor undertakes a series of CTS-Rs and evaluates the therapist's ability to deliver CBT whilst adhering to a protocol. This data is recorded. The supervisor highlights areas of the therapist's clinical work where it is thought the therapist has additional training needs. These areas might include; a therapist's ability to deliver treatment for a specific disorder, a therapist's ability to communicate effectively using written communication or a therapist's ability to deliver CBT with fidelity to the model. Where a therapist has been identified as having a specific learning need, they are directed to relevant training modules held on a bespoke e-learning platform (see following section 2.10.2,). Clear and objective learning outcomes are communicated to the therapist and these are measured at regular intervals. This methodology is only possible because of the availability of all therapy transcripts of every therapy session. Therapists are encouraged to use these transcripts to reflect on their own practice and they are offered the support and guidance of a clinical supervisory and training team to use this unique data set to enhance their practice.

2.11.2 Continuing professional development

Therapists who deliver CBT using the IECBT method have access to a unique e-learning platform which has been developed to meet their continuing professional development needs as far as is possible. The e-learning platform provides training and guidance that

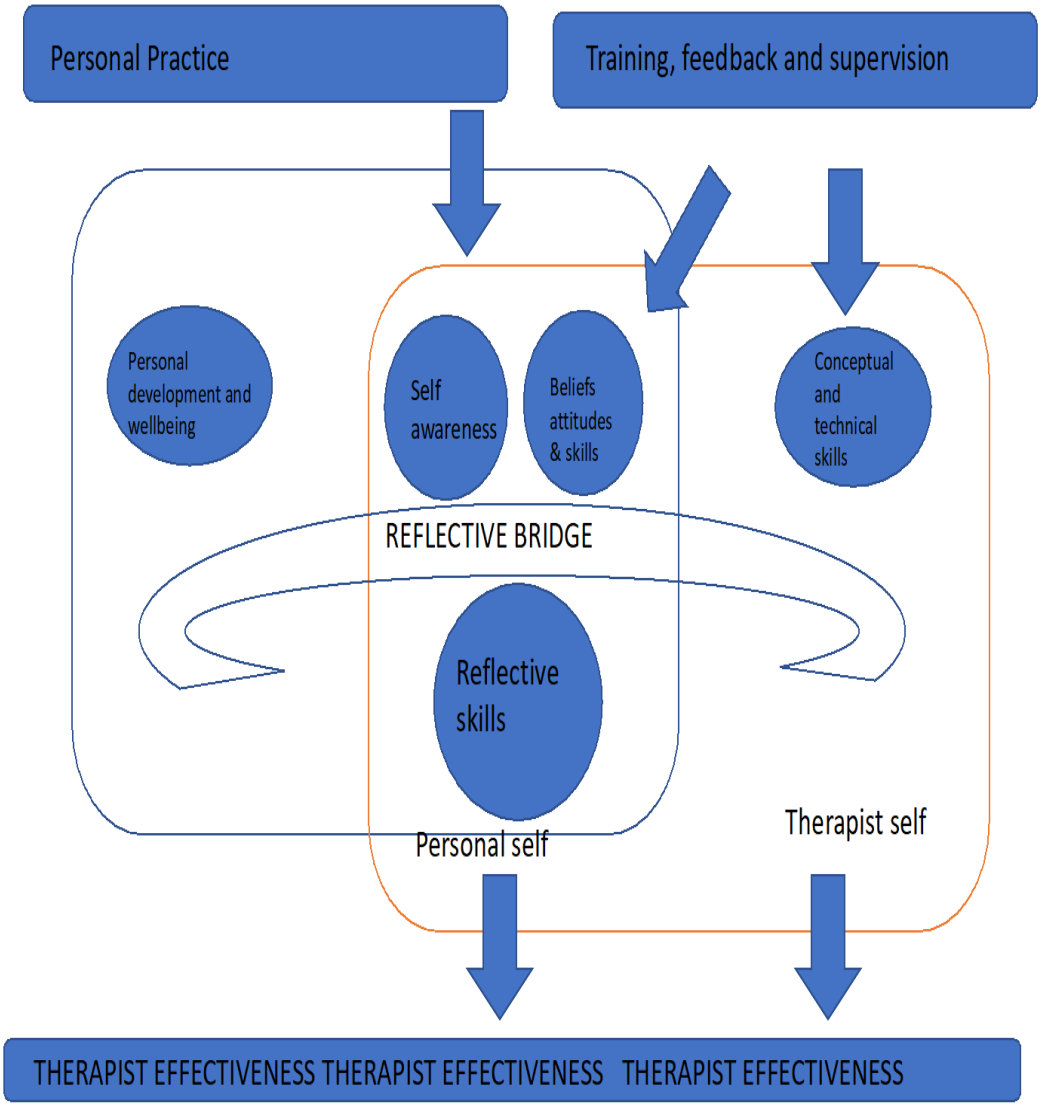
relates specifically to the IAPT training curricula (outlined earlier in this chapter). All therapists have access to the outcome data for the patients they have treated, and they have ongoing reviews of their work with the primary aim of supporting them to identify specific training needs. This e-learning platform has been developed by Ieso Digital Health, specifically to provide supervision and training to therapists working using the IECBT method. The e-learning platform provides each therapist with specific training modules to meet their individual training and supervision needs. As the platform is unique to Ieso Digital health it is shown in figure 2.6.

Figure 2.6 The IECBT e-learning platform, showing one example of a therapist's view of the modules released in order to meet their identified training needs



The underpinning conceptual context for the IECBT continuing professional development programme is a theoretical framework that emphasises the importance of self-reflection and personal practice (Bennett-Levy, 2005, Bennett-Levy and Finlay-Jones, 2018) whereby cognitive behavioural therapists are encouraged to use self-reflection in order to learn or enhance clinical skills. This personal practice model (shown in figure 2.7) regards the therapist as two interconnecting selves, the 'personal self' and the 'therapist self' (Bennett-Levy and Finlay Jones, 2018). Both selves are impacted by 4 separate domains: personal development/wellbeing, beliefs, attitudes and skills, self-awareness and conceptual/technical skills. The model asserts that where a therapist engages in personal practice, feedback, training and supervision then this directly impacts on each of the 4 domains. Bennett-Levy and Finlay-Jones (2018) emphasise the central importance of self-reflection on the learning and development process. They describe a metaphorical "reflective bridge" (see figure 2.7) between personal practice, feedback, training, supervision, the 4 personal/therapist domains and therapist effectiveness. That is to say that therapist effectiveness is enhanced by a therapist's ability to reflect on all pedagogical processes (personal practice, feedback, training and supervision).

Figure 2.7 The Bennett-Levy Personal Practice model adapted from Bennett-Levy and Finlay-Jones (2018). The model shows the interconnection between 2 aspects of the therapist, the personal self (blue rectangle) and the therapist self (orange rectangle). Both selves are impacted by 4 domains: personal development, self-awareness, beliefs/attitudes and conceptual/technical skills. The model shows how personal practice, training, feedback and supervision directly impact on a therapist's self-awareness, beliefs and skills when therapists are able to reflect on learning. This process aims to improve therapist effectiveness.

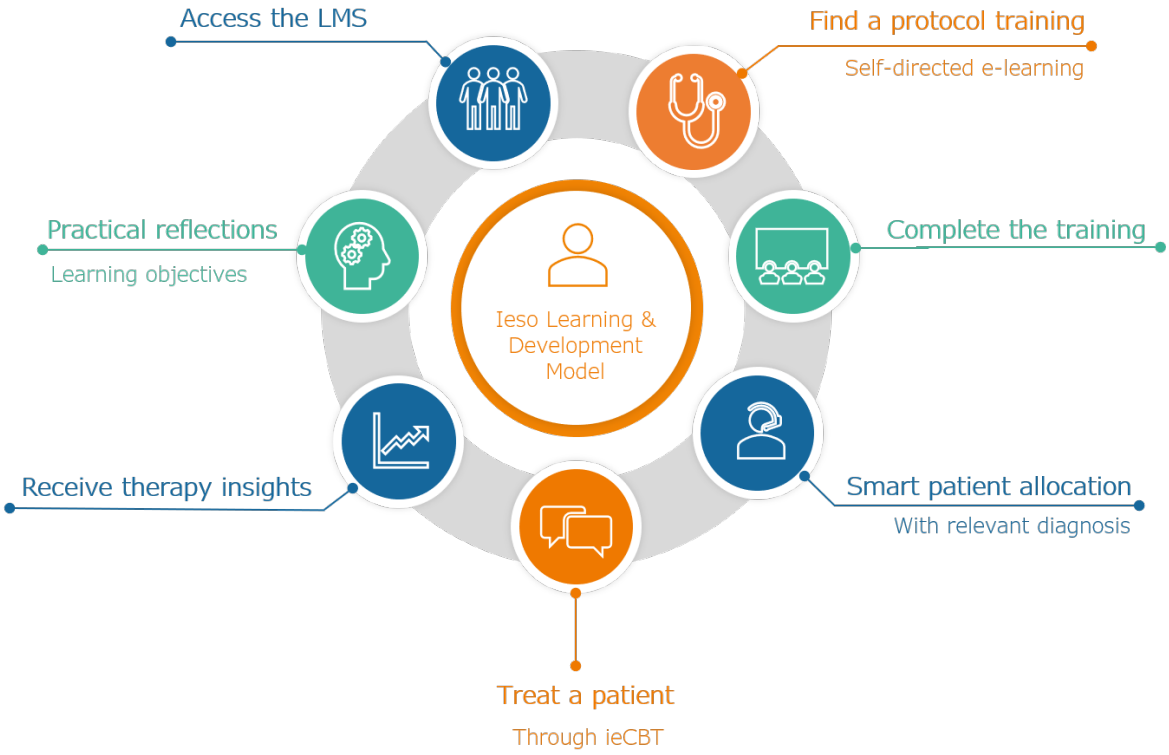


Whilst therapist self-reflection has been demonstrated to enhance practice (Bennett-Levy and Thwaites 2007) it might be argued that it is somewhat unilateral in that therapists may not be aware of their own training needs and may have 'blindspots' (Mathieson, Barnfield and Beaumont, 2008, Brosan, Reynolds and Moore, 2008). Therefore, the IECBT model includes supervisory feedback and learnings derived from clinical outcomes. Therapists are encouraged to use their own self-reflections in collaboration with the feedback and support from their supervisor and the data that has been derived from the patients they have treated. This cyclical model is shown in figure 2.8.

2.11.3 Clinical decision support tools (CDS)

In addition to continuing professional development and supervision and individual feedback based on outcome data and review of therapy transcripts, the IECBT method employs further methodology to support therapists whilst working with their patients. This includes the early development of clinical decision support tools. Just as not all therapists are able to make best use of self-reflection and may have significant blindspots regarding their practice (Mathieson, Barnfield and Beaumont) it is likely that not all therapists possess the ability to make effective clinical decisions (Barkham, Margison, Leach, Lucock, Mellor-Clark, Evans et al., 2001). This issue may significantly impact on the effective delivery psychological therapies (Borkovec, Echemendia, Ragusea and Ruiz 2001). Clinical Decision Support tools (CDS) have been described as a computerised process of providing clinicians with patient specific information, at the moment when it is required, in order to improves the quality of care that is provided (Klonoff and True, 2009).

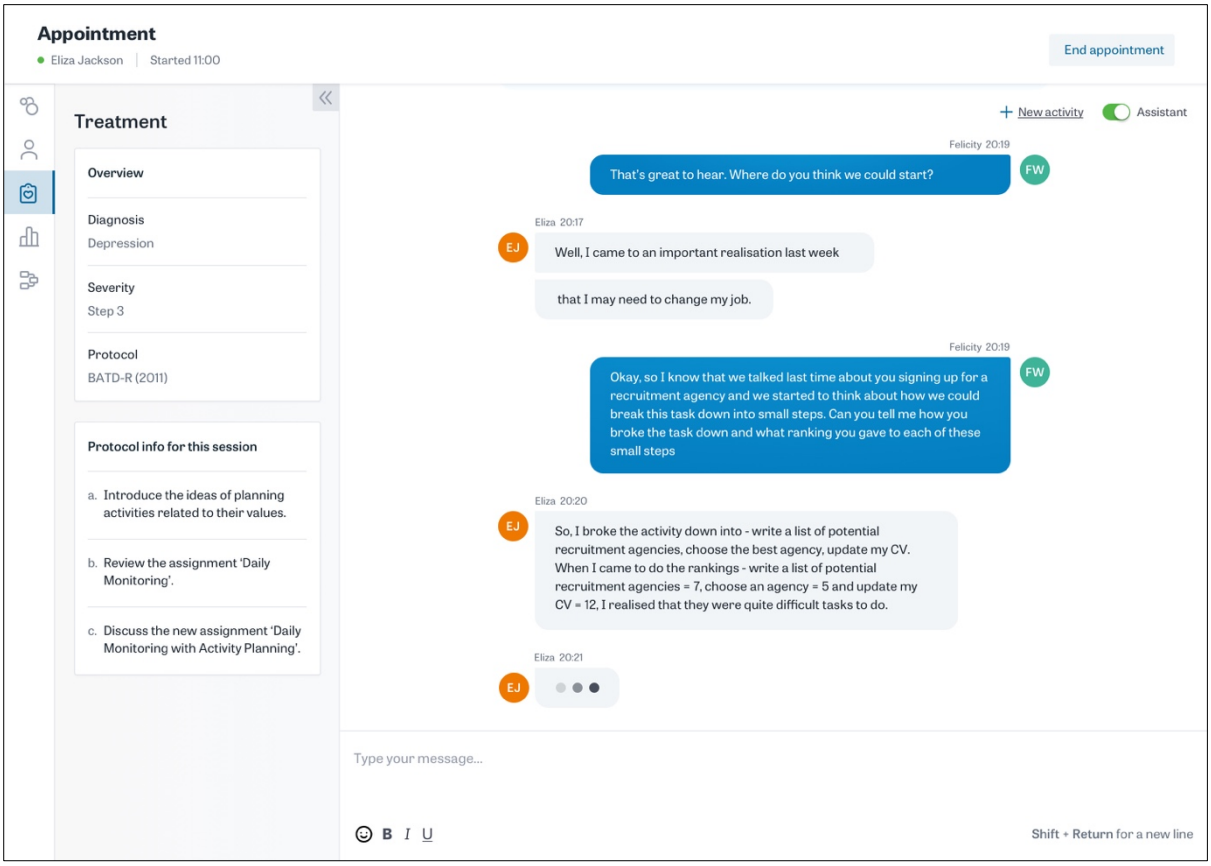
Figure 2.8 The IECBT model of continuing professional development. The figure shows how the learning management system (LMS) supports therapists to transfer theoretical knowledge into clinical practice and then receive feedback (therapy insights) on their ability to demonstrate that they have met the learning objectives whilst treating a patient.



This clinical decision support process might be defined as live digital supervision. Clinical Decision Support (CDS) tools work effectively on computer-based symptoms where data is generated or transmitted electronically (Klonoff and True, 2009). As the IECBT method is a computer-based system whereby the data (transcripts, patient outcome, patient demographics) is generated electronically it follows that CDS tools can be developed to support therapists when making clinical decisions. The CDS tools use machine learning and algorithms to make predictions about patient characteristics including likelihood of recovery, expected trajectory of recovery, possible diagnosis, severity, and likelihood of the patient dropping out of treatment (Lutz et al., 2006). The first iteration of CDS in the IECBT platform is guidance and suggestions relating to the possible diagnosis of the patient, the severity and the likelihood that the patient will drop out (Evening Standard, 2019). Figure 2.9 shows the IECBT tool on the IECBT platform.

It should be noted that it is yet to be understood to what extent therapists engage with clinical decision support tools and whether they have any impact on clinical outcomes. For example, one study suggests therapist attitudes towards the guidance they receive from clinical decision support tools (whether they like it, or agree with it) has an impact on how they engage with the guidance that has been received (Lutz, Rubel, Schiefele, Zimmermann, Bohnke and Wittman 2015). Further research and modifications will be required in order to understand these factors (Lutz et al., 2015). Clinical decision support tools are currently not used in face-to-face IAPT services.

Figure 2.9 A Clinical Decision Support tool on the IECBT platform



2.12 IECBT AND TRADITIONAL FACE-TO-FACE CBT

Section 2.10.3, earlier in this chapter, discussed the similarities in clinical outcomes between IECBT and traditional face-to-face CBT in IAPT. Both IECBT (provided by the company Ieso Digital Health) and face-to-face services deliver CBT at step 3 in IAPT for patients who present with common mental health disorders such as anxiety and depression (Layard and Clark, 2014).

2.12.1 Patient demographics

Both IECBT and face-to-face CBT services in IAPT treat patients with similar presentations. Tables 2.1, 2.2 and 2.3 show the age, reported ethnicity and clinical profile of IAPT patients respectively. The data relating to face-to-face IAPT was reported in the second round of the UK National Audit of Psychological Therapy, 2013 (Data.gov.uk., 2016) and compared to the demographic and clinical profile of patients who have had IECBT. The National Audit of Psychological Therapy (2013) reported that 65% of patients in face-to-face IAPT were female and 35% were male, compared to 69.8%, female, 28.4% male and 1.8% who declined to answer in the IECBT IAPT method.

Table 2.1 Patient age (as a %) in face-to-face CBT and IECBT

Age Group	Face-to-face %	IECBT %
18-24	13	12
25-44	44	48
45-64	35	34
65-74	6	5
75+	2	1

Table 2.2 Patient ethnicity in face-to-face CBT and IECBT

Ethnicity	Face-to-face %	IECBT %
White British	83	81
White other	5	3
Asian or Asian British	2	2
Black or black British Caribbean	2	3
Black or black British African	1	2
Other including…	7	9

Table 2.3 Reported primary presenting problem in face-to-face CBT and IECBT

Primary presenting problem	Face-to-face %	IECBT %
Depression	36	37
Mixed anxiety and depression	31	5.5
Generalised Anxiety Dis.	14	18.3
Other	14	15.9
Panic Disorder	5	4.5
Other anxiety	3	2.7
PTSD	2	3.5
OCD	2	3.6
Specific phobia	1	1.5
Social Anxiety Disorder	<1	7.5
Body Dysmorphic Disorder	<1	0

2.12.2 Waiting times to start treatment

One of the overarching aims of IAPT is to provide timely access to evidence-based psychological interventions. Therefore, the waiting time to start treatment is one of the reportable key performance indicators in IAPT (NHS England, 2018). The mean waiting time to assessment in face-to-face services in the reporting year 2017-2018 was 20.7 days (NHS Digital 2018) compared to 14 days in IECBT.

2.12.3 Therapist demographics

Both IECBT and face-to-face provision of CBT within IAPT are delivered by British Association of Behavioural and Cognitive Psychotherapy (BABCP) accredited CBT therapists. The minimum training standards required by the BABCP in order to become accredited were outlined earlier in this chapter. Therefore, it would not be expected that there was significant variance between therapists delivering CBT face-to-face and via IECBT. Furthermore, 75% of therapists delivering CBT via IECBT are also delivering CBT in face-to-face IAPT services. The most recent IAPT workforce census reports that 79% of IAPT therapists are female and 21% are male (NHS England, 2015) compared to 81.5% female therapists and 18.5% male therapists in IECBT. Of these, 88% of therapists had completed an IAPT training programme and 12% has become accredited prior to IAPT in face-to-face services compared to 73% having completed an IAPT training programme and 27% having been accredited prior to IAPT in IECBT. Table 2.4 shows the ages of therapists working in face-to-face IAPT services, taken from the most recent IAPT work force census (NHS England 2015) compared to the therapists who participated in this research.

Table 2.4 Therapist age in face-to-face CBT and IECBT (whole cohort of therapists providing IECBT)

Therapist Age	Face-to-face %	IECBT %
≤ 25	8	7
26-45	58	65
46-60	31	27
≥ 61	4	1

2.12.4 Summary

This section has examined the differences between traditional face-to-face CBT and IECBT in terms of the patients that are treated via both methods, the mean waiting times for both methods and the therapists who are working in each method. Not only are there no significant differences between therapists in both methods but 75% of the therapists working using the IECBT method for the company known as Ieso Digital Health (see www.iesohealth.com) also work part-time as face-to-face therapists in IAPT. Therefore, it is likely that the cohort of therapists, using the IECBT method, are representative of High Intensity BABCP accredited IAPT therapists.

Whilst there is some similarity in the demographic of patients treated by face-to-face therapists and IECBT therapists, there is a difference in the reported primary presenting problem. In the IECBT method only 5.5% of therapists recorded the primary presenting problem of 'mixed anxiety and depression' compared to 31% recorded by face-to-face therapists. In addition, there were significantly more patients who were treated, using the IECBT method, for the reported conditions of generalised anxiety disorder and social anxiety disorder. The smaller numbers of 'mixed anxiety and depression', in the IECBT method, might be explained by differences in continuing professional development between the two methods. Therapists, using the IECBT method, are encouraged to consider an alternative, or differential, diagnosis (most usually depression or generalised anxiety disorder) rather than

use the problem descriptor of ‘mixed anxiety and depression’. This is because there is no disorder specific protocol for mixed anxiety and depression. These operational changes to service delivery in the IECBT method are at odds with face-to-face services as IECBT has the significant advantage of having the ability to check the therapists’ clinical decision making by viewing the relevant therapy transcript. Additionally, the higher numbers of patients being treated for Social Anxiety Disorder, using the IECBT method might be explained by the online method too. It might be argued that patients with social anxiety disorder may prefer not to attend face-to-face CBT sessions because being in the social presence of another person or sitting in a waiting room with other people is anxiety provoking in itself. IECBT offers patients the opportunity to have CBT without the need to be in the physical presence of their therapist or any other person. It might therefore be argued that IECBT method affords some advantages for patients with a diagnosis of Social Anxiety Disorder and this may account for differences between patients treated using IECBT and face-to-face. There may also be differences in how therapists select a primary presenting problem between the two methods.

It should be pointed out that whilst clinical outcomes reported by both methods are roughly similar, and both patient and therapist demographics appear broadly comparable, that equivalence between the two methods has yet to be demonstrated. However, study of the descriptive statistics would suggest that there are similarities between the two methods.

The next section of this chapter will discuss the clinical outcomes reported to NHS England in more detail and will outline the problem of variance in outcomes between IAPT services and between therapists.

2.13: IAPT OUTCOME DATA: OUTLINING THE PROBLEM

Earlier sections of this chapter discussed the development of the Improving Access to Psychological Therapy programme (IAPT) and its focus on the delivery of evidence-based psychological therapies using an outcomes-based framework. Methods of delivering cognitive behavioural therapy were discussed, including face-to-face and online methods. It has been established that face-to-face CBT and IECBT are broadly similar in that they both use British Association of Behavioural and Cognitive Psychotherapy (BABCP) accredited CBT therapists to deliver therapy, therapists are required to deliver treatments that conform to National Institution of Health and Care Excellence (NICE) guidance and that the clinical outcomes from both methods are broadly the same. This section will now discuss the clinical outcomes reported to NHS England by IAPT services for patients treated at step 3. It will be argued that a problem has been identified relating to a significant variance in clinical outcomes between services and between therapists.

2.13.1 IAPT outcome-based monitoring

IAPT services are mandated to report the clinical outcomes for all patients who have had two treatment sessions or more (Clark et al., 2017). Outcomes are monitored by asking patients to complete a series of self-administered questionnaires before each therapy appointment. This process is arguably subject to human error in that face-to-face services use paper copies of the outcome measures and the therapist is required to add up the scores. Additionally, there is an element of trust that the therapist will give the patient the outcome measures before the therapy appointment and not after. It might be argued that where a patient completes the measures after the appointment, they are likely to score more highly than before the appointment. Conversely, in the IECBT methods the outcome measures are completed online (via the patient's smartphone, tablet or computer) before the therapy appointment. Unlike face-to-face CBT the therapist does not need to add up the scores as these are all calculated digitally with no requirement from the therapist. The

results from both methods are recorded on an electronic medical records system and reported to NHS England each month. The data set which holds the outcomes for every IAPT service (both face-to-face and IECBT) is held by NHS Digital (Clark et al., 2017). Most of the data is in the public domain and can be viewed online in the Public Health England's Common Mental Health Disorders Profiling Tool (Clark et al., 2017).

IAPT mandates that a service collects a minimum data set at every appointment for every patient (Clark et al., 2017). The minimum data set consists of a series of self-administered questionnaires that the patient is asked to complete before each appointment. The questionnaires include the PHQ-9 (Kroenke et al., 2001) and GAD-7 (Spitzer et al., 2006). The PHQ-9 is a nine-item, self-administered measure for depression using a 4-point Likert scale (0-3, where 0 indicates the absence of a symptom and 3 indicates greater severity). A cut off point of ≥ ten is used to indicate a diagnosis of depression in a primary care population. The PHQ-9 has been demonstrated to good reliability with a reported Cronbach's alpha of 0.89 and sensitivity and specificity of 88% (Kroenke, Spitzer and Williams, 2001). The GAD-7, is a seven-item, self-administered, measure for anxiety using a 4-point Likert scale (0-3, where 0 indicates the absence of a symptom and 3 indicates greater severity). A cut-off point of ≥ eight indicates greatest sensitivity and a clinical case in a primary care population. A cut of ≥ 15 indicates severe symptoms. The GAD-7 has been demonstrated to have good reliability with a Cronbach's alpha of 0.92, specificity of 82% and sensitivity of 89% (Spitzer, Kroenke, Williams and Lowe, 2006). In addition, therapists may use an Anxiety Disorder Specific measure (ADSM) where they believe that the PHQ-9 or GAD-7 are not sensitive to the patient's presenting problem (Clark et al., 2017). Examples of ADSMs are the Obsessive-Compulsive Inventory (Foa, Kozak, Salkovskis, Coles and Amir, 1998) sometimes used for patients with Obsessive Compulsive Disorder and the Impact of Events Scale (Weiss, 2007), used for patients who present with Post-Traumatic Stress Disorder (PTSD).

2.13.2 Definition of recovery

The IAPT definition of recovery is based on a calculation using the pre- and post-intervention scores using the PHQ-9, GAD-7 and ADSM (if used). IAPT measures clinical improvement in terms of whether a patient meets “caseness” (NHS Digital, 2016). Caseness is a term that is used to describe a patient who has sufficient symptoms to be considered a clinical case. The cut off, to be considered a clinical case (and therefore in caseness), on the PHQ-9 is a score of ≥10. The cut-off, to be considered a clinical case (and therefore in caseness), on the GAD-7 is ≥8**.** A patient is described as being recovered if they were a clinical case at the start of treatment but fell below the threshold to be considered a clinical case at the end of treatment.

2.13.3 Definition of reliable improvement

Another measure of clinical improvement is the IAPT definition of reliable improvement (Clark et al., 2017). The calculation for reliable improvement also uses the pre- and post-intervention scores of the PHQ-9 and GAD-7. A patient is said to reliably improved where they achieve anything in excess of a change of 6 points, on the global PHQ-9 score and 4 points on the global GAD-7 score (Gyani, Shafran, Layard and Clark, 2013). For example, where a patient starts treatment with a PHQ-9 score of 21 and a GAD-7 score of 18 and ends treatment with a PHQ-9 score of 14 and a GAD-7 score of 12 they would be deemed to have reliably improved. However, it should be noted that the patient in this example would not have met the definition for recovery as they had not fallen below the threshold for recovery i.e. ≤ 9 on the PHQ-9 and ≤ 7 on the GAD-7.

The IAPT definitions of recovery and reliable improvement are also shown in table 2.5, below.

Table 2.5 IAPT definitions of recovery and reliable improvement

Clinical Recovery	Reliable Improvement
Where a patient starts out in therapy as a clinical case on at least one of the measures (≥10 on PHQ-9 and ≥ 8 on GAD-7) and their post intervention scores fall out of caseness on both measures at the end of treatment.	Where a patient starts out in therapy as a clinical case (≥10 on PHQ-9 and ≥ 8 on GAD-7) and their post intervention scores fall ≥ the measurement error of the questionnaire i.e. 6 on the PHQ-9 and 4 on the GAD-7.

2.14 IAPT OUTCOMES

Whilst IAPT services are mandated to report their outcome to NHS England on a monthly basis (Clark et al., 2017) the information is reported in the public domain retrospectively. Therefore, the last reporting year available, in the public domain, is the IAPT annual report for 2017-2018, which was published in November 2018 (NHS Digital 2018). The report establishes that in this year IAPT received 1.44 million referrals, which was an increase of 3.95% from the previous year. Of these referrals, 1.01 million started treatment and 554,709 completed a course of treatment, a reduction of 2.2% from the previous year. The mean recovery rate for all services was 50.8%, an increase from 49.3%, reported the previous year. The outcomes relate to all IAPT services, face-to-face and IAPT. The report does not differentiate between the method of delivery. The mean recovery rate for IECBT for the same reporting year was 52.7%.

The increase in recovery rates to 50.8% is significant in that IAPT has demonstrated that it is possible to achieve recovery rates in excess of 50%. Layard and Clark (2014) established that national recovery rates would be a key performance indicator (KPI) when implementing the IAPT programme. Year on year, since IAPT was launched in 2008, recovery rates have incrementally improved (Clark, Canvin, Green, Layard, Pilling and Janecka, 2017). This is largely due to initiatives driven by David Clark (Clark et al., 2017, Clark, 2011) whereby learning from services with higher outcomes is disseminated to all IAPT services in an effort to improve outcomes. However, sceptics suggest that outcomes are improving because of

the creative use of data collection and statistical analyses (Binnie, 2015) whilst others (Jorm, Patten, Brugha and Mojtabai, 2017) argue that the incidence of anxiety disorders and depression are increasing rather than reducing.

2.14.1 Variance in outcomes

Whilst IAPT reports an increase in recovery rates from 49.3% in 2016-17 to 50.8% in 2017-18 there was a significant variance in outcomes between services with the lowest achieving 20.4% and the highest achieving 58.7% recovery (NHS Digital, 2018). This issue is not new, other reports on IAPT outcomes have established a similar variance in clinical outcomes between services, with one report citing the lowest recovery as 23.9% and the highest as 56.5% (Gyanni, Shafran, Layard and Clark, 2013). A later study cited more significant variance and suggested that the mean recovery rate might be as low as 23% and that if IAPT included all patients who had been referred in its analyses, the recovery rate could be as low as 12% (Scott, 2018). However, this study looked at only 90 patients who had received treatment in IAPT, all of whom had experienced some type of trauma. Not only is it possible that these patients are not representative of IAPT patients in general, but the author is concerned with the criteria for defining recovery. IAPT currently calculates recovery rates based on those patients who had a minimum of three sessions, i.e. an assessment plus two treatment sessions (Clark et al., 2017). Whereas Scott is including all patients who were referred, even those that did not have an appointment. Clearly there needs to be agreed criteria. Further critical investigation of IAPT outcomes is the focus of other studies who share the interest in variance in outcomes in IAPT (Clark et al., 2017). One study, which looked at variance in outcome from a therapist perspective, looked at 61 therapists providing treatment in IAPT (28 CBT therapists and 33 counsellors) and demonstrated that therapist recovery rates varied between 16% and 75.8% (Saxon, Firth and Barkham, 2017). The variance in recovery rates between services and between therapists is an ongoing concern. Whilst on one hand IAPT has demonstrated that it is possible to achieve its national target

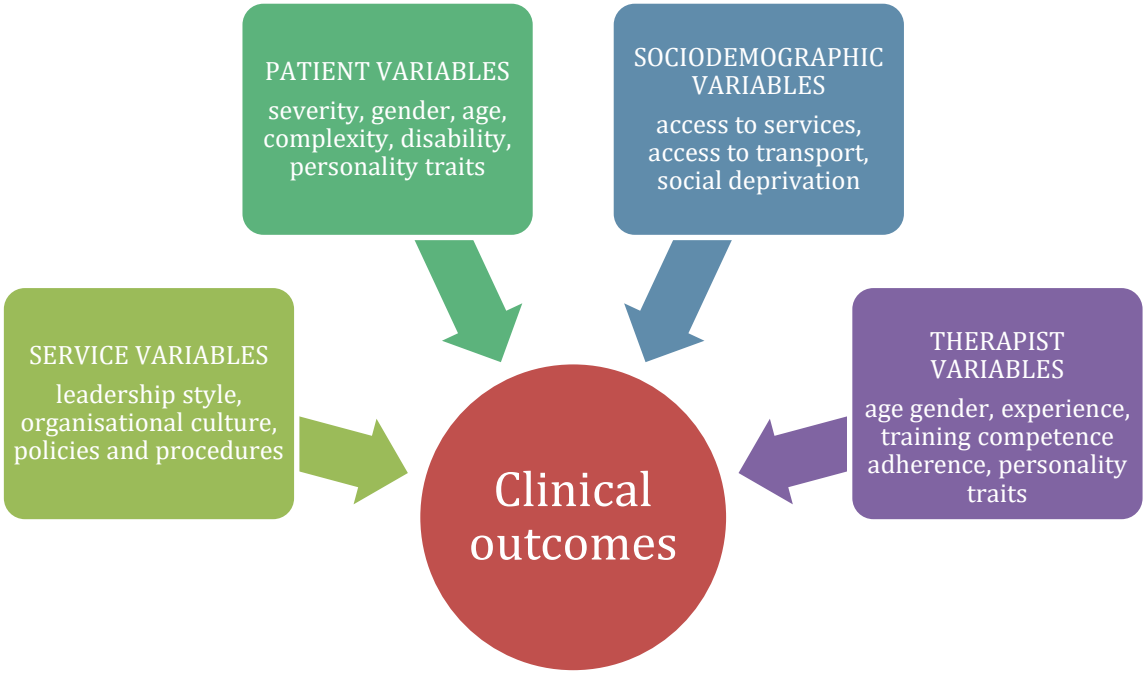
rate of 50% recovery and the trend appears to be a gradual incremental improvement (Clark et al., 2017) the variance in outcome does not appear to be changing. Furthermore, there is little more than a set of hypotheses (Clark et al., 2017) that serve to develop an understanding of why variance in outcome continues to be a problem. One of the most significant barriers to understanding variance is that it is practically impossible to assess or observe individual therapists in the course of their work with patients. Therefore, the question of why there is such a significant variance remains. It is likely that if this is not addressed it will limit future improvements in IAPT service provision and, arguably, in other CBT services too.

2.14.2 Understanding the problem

It has been clearly identified that there is a significant difference in outcome between individual IAPT services (Clark et al.,2017). Various factors have been cited as possible predictors of poor outcomes (Clark et al., 2017). Understanding why some services report better outcomes than others is key to IAPT's continued success. The key putative factors hypothesised that are likely to contribute to outcomes within IAPT have been defined as (see Figure 2.10); patient severity, patient engagement (i.e. numbers of patients who drop out of treatment and those that fail to appointments), service factors (i.e. managerial, culture of a service and processes within a service) and socio-economic factors, for example; patient income, employment status, crime and housing (Clark et al., 2017). From these factors it was suggested that patient severity did not vary across services and this was ruled out as a possible predictor of outcomes (Clark et al., 2017). Socio economic factors, on the other hand, were a predictor of outcome, although Clark (2018), has argued that some services operate in areas which include communities with high social deprivation and areas with low social deprivation and these services have reported very similar outcomes for both areas. Service factors, including patient engagement, were also reported to be predictors of outcome resulting in Clark et al., (2017) making recommendations for services to follow

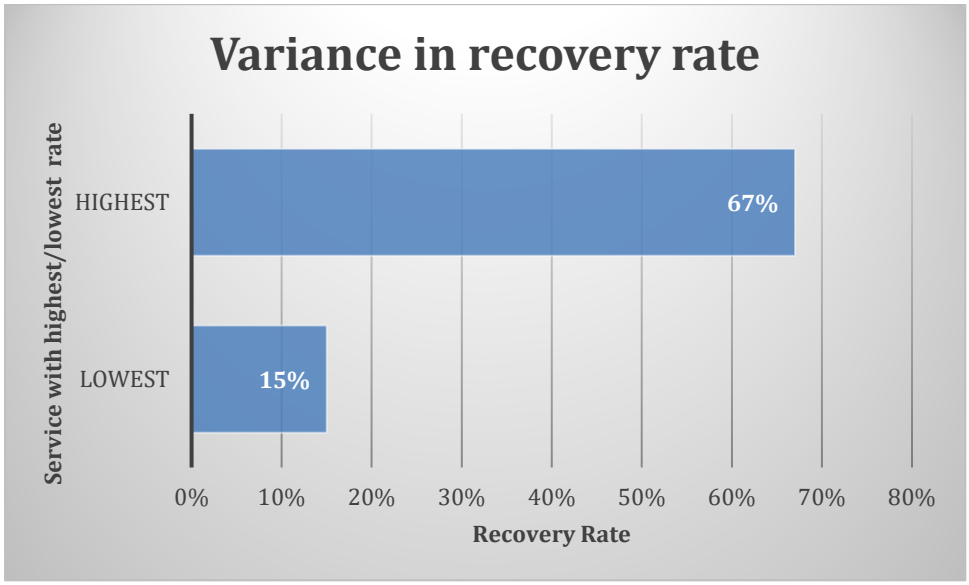
IAPT guidance by ensuring that therapy was delivered in a timely manner in accordance with the evidence base. Whilst this is sound advice, it would seem that one factor that is common to all the cited variables is the therapist. Therapists work in services and, to a certain extent, how they operate may influence waiting times, patient engagement, service culture, service processes and patient outcomes (Clark, 2018). Additionally, it might be argued that the ability of a therapist to work with someone who has greater severity or complexity (i.e. socio-economic factors or co-morbidity) should not vary from one service to another (Clark, 2018). Therefore, given that the variance in IAPT therapists' recovery rates is very similar to the variance in whole IAPT service recovery rates and IAPT therapists are situated within IAPT services it would make sense to understand what therapist variables might be influencing clinical outcomes.

Figure 2.10 Variables that have been reported to affect clinical outcome



Whilst the variance in therapist recovery rates has been established, it is unclear why this may be. Despite the transparency of data reporting cited by Clark et al., (2017), investigations into the relationship between the variables suspected of causing variance in data have yielded little more than conjecture at this stage. Clark et al., (2017) suggest that looking at service variables (the way services are run and managed) “might” be important. Despite having access to a data set that included all patients that had been treated by IAPT services in 211 Clinical Commissioning Groups (CCGs) in the years 2014-15, 2015-16, 2017-18 and 2018-19 (see figure 2.11 which shows the variance in recovery rates in 2018-19) Clark et al., have only been able to make suggestions about why such significant variance exists. Furthermore, the authors fail to identify other variables, such as therapists, as one area for further research. Given the significant investment in therapist training within IAPT, studying therapist variables may facilitate the development of training initiatives that reduce the variance in outcomes between therapists. The second United Kingdom National Audit of Psychological Therapies (Pybis, Saxon, Hill and Barkham, 2017) also report on the variance in outcomes in IAPT and suggest that further research is needed to understand the effect that therapists are having on outcomes and what therapist variables may account for this. This significant problem means that some patients may be receiving less than adequate care and is a cause for concern (Pybis et al., 2017).

Figure 2.11 Variance in recovery rates reported in 2018/19 for patients who had CBT in IAPT (n= 77,983). Source NHS Digital (2019f)



2.14.3 Literature review question

The previous sections have illustrated the problem of variance between clinical outcomes in IAPT. Having identified this as a significant problem the following question will used to review the literature: what therapist variables relate to outcome in IAPT?

2.15 CONCLUSION

This chapter has discussed the development of the IAPT programme and outlined the overarching aims, objectives and guiding principles of the programme, including the use of digital methods to deliver CBT. The use of an outcomes-based framework including the use of outcome measures and the specific calculations used to define recovery and reliable improvement have been described. Whilst it has been identified that IAPT is achieving a mean recovery rate of >50%, this chapter has identified a significant problem that relates to the variance in outcomes between IAPT services and IAPT therapists. A number of factors have been cited as possible reasons why this variability exists, these include patient variables, socio-economic variables, service variables and therapist variables. This chapter

has argued that the common variable in each of these factors is the therapist. Chapter three of this thesis will present a review of the literature in relation to therapist variables and their relationship with outcome in IAPT. The literature review will begin with an initial scoping review of the wider literature including the historical context of understanding therapist variables and their relationship with outcomes in psychological therapy across English speaking countries. This will be followed by a systematic review of the literature relating to therapist variables and their relationship with outcome in IAPT.

CHAPTER THREE: LITERATURE REVIEW

Chapter 2 of this thesis identified a problem relating to the variance in clinical outcomes between IAPT services. A number of possible causes for this variance were explored. It was argued that the common factor between each of the potential causes was the therapist, in that therapists are situated in services and should be equally qualified to treat patients with a range of complexities in the context of a step 3 IAPT service. Chapter 2 concluded with the question; what therapist variables relate to clinical outcome in IAPT? Chapter three will use this question to review the existing literature. The chapter will begin with a scoping review of the broader literature including the historical context of research in relation to variance in outcomes between therapists. This will be followed by a discussion of the wider literature drawing out themes relating to therapist variables and how they relate to outcome. The terms therapist variables and therapist effects will be defined, as they relate to the literature. This chapter will conclude with a systematic literature review focusing only on therapist variables as they relate to clinical outcomes in cognitive behavioural therapists treating patients at step 3 in the Improving Access to Psychological Therapy (IAPT) programme.

3.1 PRELIMINARY SCOPING REVIEW

The purpose of this preliminary scoping review was to examine the wider literature in relation to what is currently known about therapist variables and their effect on clinical outcome. Hart (2001) and Galvan and Galvan (2017), amongst others, argue that the function of a scoping literature review is to:

- understand how each piece of reviewed work contributes to the understanding of the researcher's subject area
- define how each piece of reviewed work relates to others
- shed light on gaps in the literature
- identify significant competing opinions

- identify what is already defined, so as to not repeat areas of research unnecessarily
- provide a sense of direction for future research;
- explicitly identify the originality of a researchers proposed area of study

This approach to "taking stock of what has gone before and identifying a niche for one's own research" (Booth, Papaioannou and Sutton 2016, p.7) is the overarching framework used for this scoping literature review. The focus of this thesis is relatively new, in that, at the time of writing (2019) the Improving Access to Psychological Therapy programme has been operating in England since 2009. Therefore, in an effort to conduct a broader search of the literature, both outside the literature on IAPT and prior to the existence of IAPT, a preliminary scoping review approach was adopted. Scoping reviews have been criticised for using less robust methodologies (Dijkers, 2009). However, a preliminary scoping review was used in this context in order to assist in developing a broad overview of the literature and help develop the most effective search strategy for a systematic review (Davis, Grey and Gould, 2009). Common criticisms of scoping reviews include a tendency to fail to disclose the inclusion criteria used for the search, failure to provide clear search terms and a failure to make explicit the search process so that it is reproducible (Dijkers, 2009). In order to mediate for these methodological concerns, but still incorporate the broader research narrative, the initial scoping review will be followed by a systematic literature review.

In this initial scoping review, English language articles were examined in relation to their relevance in addressing the following question:

1. What therapist variables correlate with clinical outcome in cognitive behavioural therapy?

This question was developed with the function of developing a better understanding of what is currently known about what therapist variables relate to outcomes in CBT, not just in IAPT but in other clinical settings and in research settings. The scoping review yielded 2,607 possible resources. A total of eight data bases were identified as most likely to contain

relevant materials. The databases identified were; SCOPUS, MEDLINE, PsycARTICLES, PsycINFO, PubMedCentral, Proquest, JOSTOR, Psychology and Behavioural Sciences. The identified data bases were searched for relevant publications between 1970 and the current time (September, 2019). The seminal papers on CBT were published in the 1970's and the most widely cited paper, on cognitive therapy, was published by Beck, Rush, Shaw, and Emery in 1979. Therefore, the years between 1970 and 2019 were selected in order to ensure that all the available literature was examined. The initial search terms used in this review were kept broad. An initial search using the Boolean operators AND, OR and the terms 'therapist effects', 'therapist variables' and outcome yielded the 2,607 possible resources in SCOPUS, 285 in MEDLINE, 121 in PsychARTICLES, 706 in PsychINFO, 1019 in PubMedCentral, 278 in Proquest, 98 in JSTOR, and 100 in Psychology and Behavioural Sciences. All English language papers and books that discussed therapist effects or variables and their relationship with clinical outcome, regardless of the clinical setting, psychological modality or the mental health disorder being treated, were examined for themes and empirical evidence. After excluding duplicates and obviously irrelevant articles, there were 227 papers that related to therapist variables and clinical outcome. In order to ensure that all available literature was examined, the reference sections from these 227 papers were examined. This process is known as snowballing and has been demonstrated as an effective approach to exploring the development of the evidence base and ensuring no studies are missed (Papaioannou, Sutton, Carroll, Booth and Wong, 2010). A further 57 papers were identified from the snowballing search. Of the resulting 284 sources 199 directly discussed therapist variables in the context of modalities other than CBT. These areas of research related to therapists working with children and young people, therapists working in other therapy modalities such as psychodynamic, systemic therapy and counselling and variables relating to clinicians working with patients with physical health conditions. Whilst these areas were beyond the scope of this study, the papers were examined to ensure that they did not include new or important material. The remaining 85 papers were examined for themes in relation to therapist variables and cognitive behavioural therapy clinical outcomes.

Thirty papers discussed cognitive behavioural therapist variables and their relationship to clinical outcomes. Eighteen papers directly discussed High Intensity cognitive behavioural therapists (or High Intensity trainees). Twelve papers discussed therapist variables in Low Intensity (step 2) clinicians. The remaining 25 sources included material where it was unclear or unspecified what type of therapy the therapists were delivering or the therapy that was described was guided-self-help.

3.1.1 Scoping Review Theme 1: The historical development of understanding therapist variables

The findings from the preliminary scoping review will commence with a discussion relating to the historical development of research in the area of understanding variance between therapists. This section will then go to discuss the other themes that were found in the scoping review.

The term cognitive behavioural therapy (CBT) was first used in the scientific literature in the early 1970's (see Beck, 1970, Shealy,1972, Bandura, Adams and Beyer, 1977). The first paper discussing variance in outcome between psychological therapists appears to be Ricks (1974) who described two therapists, one of whom had very poor outcomes and the other outstanding outcomes. Allegedly the outstanding therapist's patients called him "*supershrink*." This seminal paper outlined the impact therapists can have on patients. A later paper (Martindale, 1978) highlighted the variance between therapists and suggested that 63% of efficacy trials ignored the impact that individual therapists could have on outcomes. So much so that he suggested that it was likely the results from trials may not be generalisable because of this. The first meta-analysis in this research area (Crits-Christoph, Baranackie, Kurcia, Beck, Carroll, Perry et al., 1991) commented on possible therapist variables that may be associated with variance between therapists including therapist experience and the use of a treatment manual. However, major limitations of the research in this field were acknowledged including small sample size and a concern about the lack of

real-world, naturalistic studies. Two real-world studies (Okiishi, Lambert, Nielsen and Ogles, 2003 and Brown, Lambert, Jones and Minami, 2005) found significant variance between therapists in their studies, noting that patients allocated to therapists with the highest outcomes showed a threefold higher recovery rate than other patients. Later meta-analyses (Baldwin and Imel, 2013) again comment on the limitations of studies suggesting that research in this field should be specifically designed as therapist effects studies rather than repurposing data from trials. This supports Martindale's earlier argument that variance between therapists in efficacy trials is not generalizable because this type of research is designed to minimize variance and maximise therapist efficacy. Therapist effects research has developed significantly, adopting robust methodological and statistical methods usually involving a statistical method called multilevel modelling. This statistical method originated in educational research and is used when data is said to be nested (Barkham, Lutz, Lambert and Saxon, 2017). In education, students are nested in classes and classes are nested within schools. In psychological therapy patients are nested within therapists and therapists may be nested within services. Whilst these types of studies have added a great deal to the literature on variance in outcomes other methods, most usually process outcome research, have also contributed to knowledge. The themes from both these research approaches, are discussed below.

3.1.2 Theme 2: Variance in outcome and therapist competence

One of the predominant themes in the literature relating to therapist variables and their relationship with clinical outcome is the theme of therapist competence as it relates to variance in clinical outcomes. As the evidence for CBT became widely established in the 1980s and 1990s, a problem emerges in the literature. This problem relates to significant variance in clinical outcomes between research trials and real-world clinical settings (DeRubeis and Feeley 1990, Chambless and Ollendick, 2001 and Trepka, Rees, Shapiro, Hardy and Barkham, 2004). It is questionable whether the results from clinical trials are

generalizable in clinical practice (Lowe, Bunnell, Neeren, Chernyak and Greberman, 2011). One of the factors that may contribute to this disparity is the differences in therapist competence between the two settings (Whisman,1993). Therapist competence has been variously defined as a therapist's ability to deliver a treatment to an acceptable standard (Fairburn and Cooper, 2011), therapist skillfulness (Kazantis,2003) and a therapist trait (Mansell, 2008). It is noteworthy that some authors fail to define competence (see Kuyken and Tsivrikos, 2009 and Keen and Freeston, 2008), or merely state that competence is a score on the CTS or CTS-R (Schmidt, Strunk, DeRubeis, Conklin and Braun, 2018). This latter approach belies an acceptance (by these authors) that the individual CTS or CTS-R items are the skills that define competence. Whilst there is clearly some difficulty in defining the term 'competence', the issue that therapists' competence might vary between research settings and clinical settings is widely documented (Whisman,1993). Unlike clinical settings the majority of research trials select therapists based on a competence-criteria (Whisman,1993) thus ensuring that therapists delivering treatments in an intervention arm of a trial have a high level of competence. The first measure of therapist competence was the Cognitive Therapy Scale (Young and Beck, 1988). The Cognitive Therapy Scale (CTS) was developed to measure therapist fidelity to the CBT model in the United States of America's National Institute of Mental Health Treatment of Depression Collaborative Research Program (Elkin et al., 1989). This randomized controlled trial is one of many that investigate whether one form of psychological therapy is superior to another. In this case Elkin et al., (1989) investigated the efficacy of CBT and Interpersonal Psychotherapy in 250 patients who met the diagnostic criteria for major depression disorder. Patients were randomly assigned to one of four treatment groups: CBT, Interpersonal Psychotherapy, antidepressant treatment (Imipramine Hydrochloride) or a placebo (patients were given a tablet which contained no active ingredients). Importantly, the findings were inconclusive due to a disparity between the severity of symptoms between the four patient groups. Whilst the study had taken care to mediate for variability amongst therapists' they may not have considered variability between patients. Ten years after this study, a post hoc analysis of

the data (Shaw et al.,1999) showed a weak relationship between higher scores on the CTS and the Hamilton Rating Scale for Depression (Hamilton, 1960). They reached the conclusion that, because of the weak findings, the CTS was not the right tool to assess therapist competence. However, this might not have been the right conclusion. The Ellkin et al., (1989) trial had selected therapists who they deemed to be more competent and therefore the cohort of therapists may not have been representative of the general population of therapists (Whisman, 1993). To reject the CTS on the basis of these findings may have been pre-emptive as later studies, in clinical settings, report some correlation between therapists and trainees therapists who had higher scores on the CTS and patient outcomes (Kingdon, Tryer, Seivewight, Furguson and Murphy, 1996 and Milne, Baker, Blackburn, James and Reichelt, 1999). However, the results from both of these of studies have been criticised as methodologically flawed (Trepka, Rees, Shapiro, Hardy and Barkham, 2004). One study (Kingdon et al., 1996) failed to investigate the strength of the relationship between competence and outcome. Unfortunately, the Milne et al., (1999) longitudinal study of CBT trainees failed to control for the progress trainees made over the duration of their training and, therefore, their finding that competence correlated positively to clinical outcome may have only been as a result of the students' training. Drawing on the outcomes of these earlier studies Trepka et al., (2004), went on to assess the competence of 6 British psychologists using the CTS. The 6 therapists treated a total of 30 patients who met the diagnostic criteria for depression. The therapists were trained to use the Beck, Rush, Shaw and Emery (1979) Cognitive Therapy protocol for depression. All the treatment sessions were recorded via audio tape and one recording was selected, at random, to be assessed using the CTS. The resulting CTS scores were compared to clinical outcome using the Beck Depression Inventory (Beck, Ward, Mendelson, Mock and Erbaugh, 1961). Trepka et al., (2004) reported that that was a weak correlation ($r = -.28$) between clinical outcome and the CTS. They suggested that it is only possible to detect a relationship between competence and outcome when there is a variance in competence between the therapists being investigated.

The CTS was widely used in the United Kingdom until 2002. The original CTS (Young and Beck, 1980) has been superseded by a revised version (Young and Beck 1988) and this is still a commonly used tool used to analyse the relationship between therapist competence and clinical outcome in the United States of America. Whilst the CTS is a widely used tool to assess therapist competence, it is important to note that there is some disagreement about whether the tool is fit for purpose. Blackburn, James, Milne, Baker, Standart, Garland and Reichelt (2001) argue that that the CTS does not enable raters to clearly differentiate between levels of competence. Whisman (1993) states that the CTS scoring system makes it difficult for raters to be objective. Furthermore, it has been argued that the CTS does not adequately assess all the domains necessary to be a competent therapist and that there is significant overlap between the items (Shaw et al., 1999). These 'design' flaws in the original CTS led to the development of the Cognitive Therapy Scale- Revised (Blackburn et al., 2001) which is commonly used (see chapter 2 section 2.5) in the UK. Despite its wide adoption there is still little evidence that there is a strong relationship between higher CTS-R scores and clinical outcome (Kazantis, Clayton, Cronin, Farchione, Limburg and Dobson 2018). This uncertainty about whether the CTS or CTS-R effectively measures therapist competence has fueled attempts to develop new tools (see Muse and Macmanus 2016) but, these have not been widely adopted to date and are yet to demonstrate a stronger relationship with clinical outcome than the CTS or CTS-R. Furthermore, there appears to be an intersect between therapist competence and therapist adherence, to a treatment protocol whereby it might be hard to measure one without the other (Kazantis, 2003). Therapist adherence is discussed as Theme 3, below.

3.1.3 Theme 3: Identifying and measuring therapist adherence

Therapist adherence has been described as one of the specific factors integral to the delivery of a psychological therapy, such as CBT (Wampold, 2015). That is to say, cognitive behavioural therapists are taught to deliver CBT according to evidence-based treatment

protocols (see Chapter 2 for a more detailed discussion of evidence-based protocols). Therefore, adherence is defined as the delivery of theoretically orientated mechanisms of change, that are specific to a method of treatment, such as CBT (Farmer, Mitchell, Parker-Guilbert and Galovski, 2017). It is assumed that therapist adherence to these protocols has a strong causal relationship to clinical outcome (Layard and Clark, 2014, Clark, 2011, Clark 2018). However, there are relatively few studies that examine the relationship between therapist adherence and clinical outcome (Webb, DeRubeis and Barber, 2010). Whilst there is an assumption that the recovery rates achieved in the CBT research trials are due to therapist adherence to a treatment protocol (Waller 2009), it should be noted that researchers pay strict attention to therapist selection, training, monitoring and supervision in order to increase the likelihood that the therapists will adhere to the prescribed protocol (Roth, Pilling and Turner, 2010). This emphasis on the importance of adherence in CBT research trials corroborates the argument that adherence to a protocol may have a direct relationship with outcome. However, very little is known about the devices (e.g. additional training, supervision, therapist selection) that researchers use to maximise therapist adherence in clinical trials (Roth, Pilling and Turner, 2010). Therefore, it is difficult to replicate the findings in clinical practice. However, it might be argued that even if researchers were explicit about the devices they used, within their research, that it might prove impossible for clinical services to replicate them because of lack of resources. Perhaps, not surprisingly, most of the studies relating to therapist adherence and clinical outcome are undertaken in the context of a clinical trial whereby the researchers' rate to what extent the therapist adhered to the treatment protocol that is under investigation. Clearly, it is important for the therapist to adhere to the protocol in these studies if the protocol in question is to be satisfactorily evaluated (Roth, Pilling and Turner, 2010). Methods for monitoring therapist adherence vary but arguably monitoring adherence requires access to recordings of whole therapy sessions for the entire treatment episode (i.e. every treatment session). Given that a treatment protocol is delivered sequentially at every session it would seem unrealistic to assess adherence using less than every session.

However, studies that rate adherence tend to use between 1 and 3 sessions (Weck et al., 2016,) and then proceed to report that there is a strong relationship between adherence and outcome on this basis (Weck et al.,2016 and Youn, Xiao,Kim, Castonguay, McAleavey, Newman and Safran, 2017 and Resko, Walton, Chermack, Blow and Cunningham, 2012). One study (Ehlers, Grey, Wild, Stott, Liness, Deale et al., 2013) did not review any therapy recordings and based their rating of adherence from reviewing therapists' clinical notes. Based on these studies and the related trials, the importance of adherence to a treatment protocol is strongly emphasized in the literature. Indeed, adherence to a protocol is a central element of the curriculum for CBT therapists in the United Kingdom's Improving Access to Psychological Therapy (IAPT) post-graduate clinical training programme (see Chapter 2). It has been inferred that therapists who fail to adhere to the treatment protocol have lower outcome rates (Waller 2009, Waller and Turner 2016 and Gyani et al., 2013). However, it should be noted that this guidance is based on trials that, on the whole, have little or no access to therapy recordings. Furthermore, there are very few real-world studies that examine the relationship between therapist adherence and clinical outcome. Of those that do, none examine whole treatment episodes (every treatment session) and, again, some make claims without ever reviewing therapists' work (see Clark, 2014, Waller, Stringer and Meyer, 2012, Liness, Lea, Nestler, Parker and Clark, 2017). For example, in Clark's (2011) paper where he reviews the treatment of "over 3500 patients" (Clark, 2011 p.320) and Gyani, Shafran, Layard and Clark, (2013) where they reviewed the treatment of 19,395 patients there is an explicit argument that adherence to National Institute of Health and Care Excellence (NICE) guidelines (which recommend disorder specific protocols) was strongly correlated with clinical outcomes. However, both papers admit that there was no evidence to support this correlation other than therapists self-reporting that they were following the guidelines. One issue with self-reporting is that it is open to subjectivity (Mathieson, Barnfield and Beaumont, Waller, 2009, Brosan, Reynolds and Moore 2008, Waller and Turner, 2016) and therapists may have a tendency to overrate themselves. It might be argued that this effect is amplified when IAPT therapists are interviewed about their practice in a large

research study conducted by one of the founders of IAPT. It would seem less likely that a therapist would present an honest reflection of their clinical practice when IAPT has an explicit expectation that a therapist will use a disorder specific protocol. In order to understand whether there is a relationship between adherence and outcome in real-world settings it will be necessary to conduct research where access to recordings or observations of whole episodes of treatment at sufficient scale (Webb, DeRubeis and Barber 2010).

3.1.4 Theme 4: Identifying and measuring the therapeutic alliance

The most commonly researched variable, in the psychological therapy process outcome research, is the therapeutic relationship. A therapists' ability to develop and maintain a therapeutic relationship has been frequently cited as essential to achieving good clinical outcomes (Trepka, Rees, Shapiro, Hardy and Barkham (2004). Rogers (1957) asserted that developing and maintaining a therapeutic relationship, using empathy, positive regard and sincerity, was sufficient to achieve good clinical outcomes. Goodyear, Wampold, Tracey and Lichtenberg (2017) point out that most of the literature relating to the importance of the therapeutic relationship is based on the work of Karl Rogers. They argue that despite Rogers' strong assertions there is no evidence to support his claims.

The therapeutic relationship is a polytheoretical phenomenon that is used to describe the connection between a patient and a therapist. The qualities of the connection include, collaboration, shared goals and a shared understanding of the problems being addressed in therapy (Bordin,1979, Horvath and Luborsky, 1993, Horvath, 2006). Rogers (1957) included the conditions of "unconditional positive regard" (p.208) and "empathy and genuineness" (p.210) in his conceptual framework of the therapeutic relationship which he believed was the causal factor for therapeutic change. The subject of understanding therapeutic processes and how these relate to outcome has been explored in just over 200 studies. Meta-analyses suggest that there is a correlation between the therapeutic relationship and

clinical outcome (Horvath, Del Re, Fluckiger and Symonds, 2011, Fluckiger, Del Re, Wampold, Symonds and Horvath, 2012). Whilst process-outcome research has repeatedly suggested that there is a strong relationship between a therapists' ability to develop and maintain a therapeutic relationship and therapy outcomes, it would appear that this is the case regardless of the modality of treatment being delivered (Llewelyn and Hardy, 2001). Frank and Frank (1991) and the later work of Wampold (2015) define the therapeutic relationship as a 'common factor' that is necessary in all modalities of psychotherapy. However, whilst it is agreed that a strong therapeutic alliance is necessary in order to achieve good clinical outcomes, it is also acknowledged that other, specific, factors are also important (Wampold, 2015, Wampold, Baldwin, Holtforth and Imel, 2017, Hill and Castonguay, 2017). Further research is required to understand how specific factors, such as therapist competence and adherence to specific treatment protocols, relate to clinical outcome (Hill and Castonguay, 2017).

3.1.5 Theme 5: Therapist demographics

Section 3.1.4 reported that the therapeutic alliance and its relationship with clinical outcome, regardless of the psychotherapy modality, is the most commonly occurring theme in the literature. In fact, it has been suggested that no further research is required into the specific effects of the therapeutic relationship as there is already sufficient literature (Baldwin and Imel 2012). It has been argued that other variables such as therapist age, gender, years of experience and academic achievement have also been thoroughly investigated and no new learning is being derived (Baldwin and Imel 2012, Wampold, Baldwin, Grosse Holtforth and Imel, 2017). A review of the literature in the fifth edition of Bergin and Garfield's Handbook of Psychotherapy and Behaviour Change (2004) by Beutler, Malik, Aliomohamed, Harwood, Talebi, Noble and Wong (2004) suggested that a therapist's gender, age, years of experience or academic background were unrelated to clinical outcome. Later studies have

confirmed these findings with Anderson, Ogles, Patterson, Lambert and Vermeersch, (2009), Schottke, Fluckiger, Goldberg, Eversmann and Lange, (2016) and Chow, Miller, Seidel, Kane, Thornton and Andrews, (2015), all confirming that therapist gender had no significant relationship with outcome. Similar findings are reported confirming that therapist age has no relationship with outcome (Chow et al., 2015, Wampold and Brown, 2005). More controversially, years of experience and therapist academic or clinical qualification have also been repeatedly confirmed to have no relationship with clinical outcome (Tracy, Wampold, Lichtenberg and Goodyear, 2014, Chow et al., 2015, Kraus, Bentley, Boswell, Constantino, Baxter and Castonguay, 2016, Wampold and Brown, 2005). The strength of the evidence and a lack of an opposing argument (to date) leads to the conclusion that these variables warrant no further comment or investigation. Furthermore, as this position becomes more accepted there is a move towards other types of study which move away from fixed effects such as therapist age and years of experience and a move towards random effects of therapists on clinical outcome (Baldwin and Imel, 2012). Whilst there is some lack of agreement about the exact definition between fixed effects and random effects in the statistical literature (Martindale, 1978, Gelman, 2004)**,** random effects, in this context, have been defined as the combined effects of all therapist variables on patient outcome (Baldwin and Imel, 2012). These studies (known as therapist effects studies) are becoming more common and are discussed in Theme 6, below.

3.1.6 Theme 6: Therapist Effects

Therapist effects differ from therapist variables in that therapist effects look at the combined effects of all therapist variables (Hill and Castonguay, 2017). That being said, therapist effects can only be measured where there is some variance between therapists and some therapists are achieving better outcomes than others. Variance between therapists is widely reported in research settings and real-world settings (Saxon and Barkham, 2012 and Nissen-Lie, Goldberg, Hoyt, Falkenstrom, Holmqvist and Nielsen 2016). The growing body

of therapist effects literature suggests that the effect an individual therapist can have on their patient is critical to patient outcome (Barkham, Lutz, Lambert and Saxon, 2017). Therefore, where a patient is allocated to a less able therapist, they are significantly less likely to recover (Nissen-Lie et al., 2016). For example, one study (see Saxon and Barkham, 2012) found that out of 119 therapists studied, 19 therapists had consistently worse clinical outcomes than the other therapists and the patients treated by these therapists were less likely to recover. Meta-analyses of therapist effects studies suggest that between 4% and 10% of variance in outcome is due to therapist effects (Webb, De Rubeis and Barber, 2010, Baldwin and Imel, 2013 and Wampold and Imel 2015, Johns, Barkham, Kellett and Saxon, 2018). However, significant methodological issues, relating to the variation in research design, outcome measures and variables have been reported (Johns et al., 2018). Additionally, issues are reported regarding small sample sizes in all the meta analyses. Arguably, these methodological issues hamper progress in this important research area. Undoubtedly, these issues prompted Schiefele, Lutz, Barkham, Rubel, Saxon, Schulte et al., (2016) to provide helpful guidance on sample sizes in therapist effects studies. The authors prompt future researchers to increase sample sizes.

The methodological issues found in the current literature make it difficult to disentangle the various studies and arguably make it harder to draw conclusions (Johns et al., 2018). Regardless of these issues, the body of literature examining therapist effects adds a great deal to the global understanding of how therapists' effect clinical outcome. However, many questions remain. Perhaps the most important include, which therapist variables have the greatest impact on outcome and how can therapist effects be reduced, whilst also improving recovery rates. One limitation of the therapist effects literature is that most therapist effects research is conducted on raw data with little, or no, access to live therapy recordings (Hill and Castonguay, 2017). This is the same issue that relates to the literature on therapist competence and therapist adherence. There is general agreement that further research should include correlational process studies using recordings of live therapy sessions alongside outcome data (Hill and Castonguay, 2017, Barkham et al, 2017).

3.1.7 Summary of findings from the preliminary scoping review

The inherent difficulties associated with conducting research, with the aim of understanding what therapist variables are associated with clinical outcome, is widely discussed in the available literature (Wampold et al., 2017). There is even some argument that the question is not researchable (Fiske, 1977). However, the aim to understand the characteristics and behaviours of effective therapists in order to improve outcomes for patients has been the main motivating factor for researchers in this field. Early attempts to study therapist variables (Ricks, 1974 and Martindale, 1978) highlighted the importance of understanding therapist variability but also brought attention to the problems that are encountered in adding to the knowledge. Most commonly, sample sizes are small for both therapists and the patients they have treated (Maas and Hox 2004). Table 3.1 shows the sample sizes for the major studies conducted to date. In addition to the issue of small sample sizes many of the studies are limited by their use of therapist surveys (asking therapists what they do with their patients, see Clark, 2011 and Liness, Lea, Nestler, Parker and Clark, 2016) or allowing therapists to select their own therapy tapes for rating (Wampold et al., 2017). Furthermore, very few studies are able to review the entire work of all the therapists that are studied and, where this is possible, the researchers have only been able to use a very small sample size (see Jeong Youn, Xiao, Kim, Castonguay, McAleavey, Newman and Safran, 2017). Other concerns have been the difficulty in disentangling studies when they frequently use different patient outcome measures, different tools and methods to assess therapist competence or adherence and use therapists drawn from different modalities (Wampold et al., 2017). However, this preliminary scoping review has shown that there is general agreeance that the therapeutic alliance is strongly related to outcome but only as a common factor or fixed effect and that other specific factors are required to achieve good clinical outcome.

Table 3.1 Sample sizes of the major studies

Article	Author/s and year	Patient	Therapist	Study Type
1	Kazantis et al, (2018)	50	4	Correlation study
2	Youn et al., (2017)	3	3	Observational study
3	Kuyken & Tsivrikos (2009)	69	18	Observational study
4	Nissen-Lie et al., (2016)	520	30	Therapist effects study
5	Weck et al., (2016)	84	34	Correlation study
6	Weck et al., (2015)	68	26	Correlation study
7	Ehlers et al., (2014)	121	6	Randomised controlled trial
8	Branson et al., (2015)	1247	43	Naturalistic observational cohort study (trainees)
9	Brosan et al., (2006)	24	24	Naturalistic observational
10	Webb et al., (2012)	105	6	Process-outcome study
11	Saxon & Barkham (2012)	10786	119	Therapist effects
12	Trepka et al., (2004)	30	6	Correlation study
13	Haug et al., (2016)	82	22	Correlation study
14	Goldberg et al., (2016)	6591	170	Naturalistic longitudinal study
15	Resko et al., (2012)	60	6	Construct validity study
16	Ginzburg et al., (2012)	38	10	Therapist effects and correlation study
17	Brown et al., (2013)	176	14	Randomised controlled study
18	Farmer et al., (2017)	68	8	Process-outcome study
19	Laska et al., (2013)	192	25	Therapist effects study
20	Boswell et al., (2013)	226	21	Correlation study
21	Shaw et al., (1999)	40	8	Correlation study
22	Horvath et al., (2011)	-	-	Meta-analysis
23	Chow et al., (2015)	1632	17	Therapist effects study
24	Keen & Freeston (2008)	52	5	Naturalistic observational cohort study (CBT trainees)
25	Llewelyn & Hardy (2001)	-	-	Literature review
26	Gyanni et al., (2013)	“up to” 19395	Not stated	Naturalistic observational cohort study
27	Tracy & Wampold (2014)	-	-	Literature review
28	Kingdon et al., (1996)	70	30	Correlation study
29	Schmidt et al., (2018)	6	6	Pilot study
30	Ricks (1974)	Not stated	2	Case study
31	Elkin et al.,	119	17	Therapist effects

Additionally, meta-analyses of therapist effects studies have estimated that between 3% and 8% of variance in outcome is due to random therapist effects. Random therapist effects have been defined as the effect of all therapist variables (e.g. therapist demographics, competence, adherence, ability to maintain a therapeutic relationship and other therapist qualities). Whilst the impact of therapist variance on patient outcomes is of significant concern, and this has motivated research in this area, there are still several unanswered questions in relation to what therapist variables are related to outcome and how might this knowledge be used to improve outcomes for patients. Improving outcomes for patients is one of the aims of the Improving Access to Psychological Therapy (IAPT) programme and, whilst the mean recovery rates for the programme are improving year on year (Clark, 2018), there is a significant variance in outcome with some services reporting recovery rates as low as 8% and others 80%. This is an ongoing concern and the focus of a small but expanding body of research. The following section will discuss a systematic review of the literature relating to therapist variables and their relationship to outcome in High Intensity CBT (step 3) IAPT.

3.2 SYSTEMATIC LITERATURE REVIEW

Having conducted a preliminary scoping review of the wider literature, the purpose of this systematic literature review was to identify which therapist variables have a statistically significant relationship with clinical outcome specifically in High Intensity cognitive behavioural therapists, treating patients at step 3, in IAPT.

3.2.1 Objectives

The objectives for this systematic literature review are:

- To search for primary articles that investigate the relationship between therapist variables and clinical outcome in High Intensity CBT therapists treating patients at step 3 in IAPT.
- To compare primary articles on therapist variables and their relationship to outcome in High Intensity CBT therapists treating patients at step 3 in IAPT
- To identify the research designs, statistical analyses and limitations of the studies that have been undertaken
- To identify recommendations for further research

3.2.2 Methodology

This systematic literature review was conducted following the Preferred Reporting Items for Systematic Reviews and Meta Analyses (PRISMA, 2009) checklist and guidelines, with a particular focus on establishing explicit objectives, clearly identifying search terms and inclusion criteria as well as a diagrammatic representation of the literature reviewed (PRISMA diagram see figure 3.1). An a priori protocol was developed to guide this systematic review establishing a clear structure, process and framework. Liberati et al., (2009) argue that it is essential to transparently report the processes involved when conducting a review of the available literature in order that it is generalizable. The PRISMA guidelines were developed to provide a robust structure for researchers to adopt in order to enhance the quality of a systematic literature review (Booth, Sutton and Popaioannou, 2016). The review of the wider available literature, relating to how High Intensity cognitive behavioural therapist variables relate with clinical outcome in IAPT, was conducted by methodically searching for all available materials and then appraising and synthesising the materials and available outcomes. All materials reviewed in this search were assessed in relation to their quality and relevance to the research question, using a 0 to 6 rating scale, where 0 represented poor-quality evidence and 6 represented very high-quality evidence (Centre for Reviews and Dissemination, 2008). The articles were examined in relation to

their relevance in addressing the following question:

1. What therapist variables relate to clinical outcome in High Intensity Cognitive Behavioural Therapists working in the Improving Access to Psychological Therapy (IAPT) programme?

This question was developed with the function of developing a better understanding of what is currently known about High Intensity cognitive behavioral therapist variables and how these might relate to clinical outcomes in IAPT. Following initial screening, articles were included if they met the following inclusion criteria:

i) Published between 2009 and the current time (July 2019)

ii) Published in the English language

iii) Research related to qualified High Intensity (step 3) cognitive behavioural therapists working in IAPT

iv) Results included statistical findings on the relationship between therapist variables and IAPT definitions of recovery and reliable improvement

v) Primary source

vi) Patients were adults ≥ 18 years of age

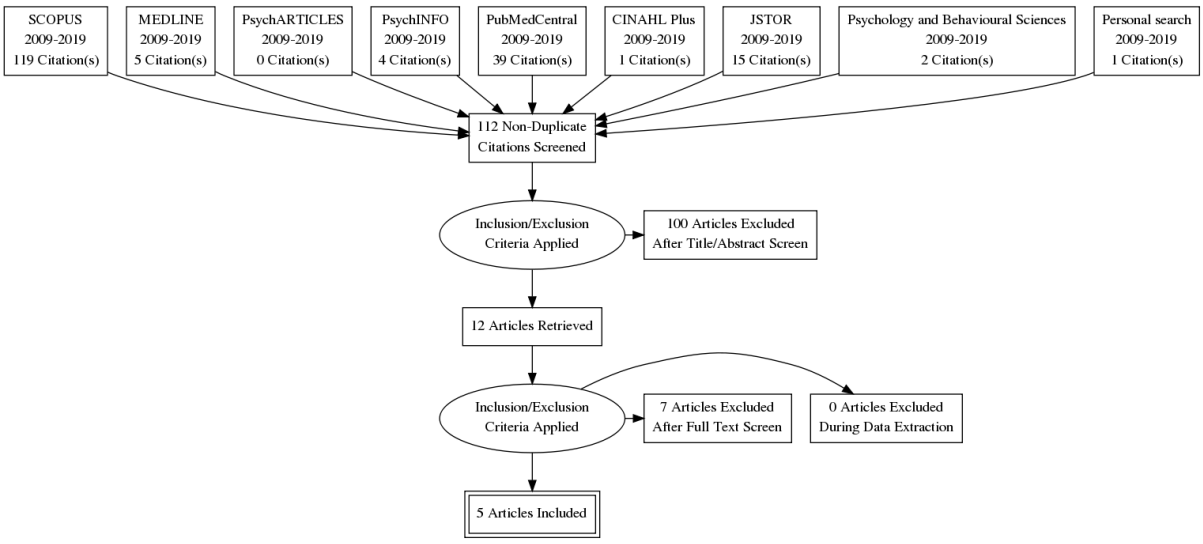
The search included articles published between 2009 and the current time (July, 2019). The Improving Access to Psychological Therapy Programme (IAPT) began in 2009 and therefore the years between 2009 and 2019 were selected in order to ensure that all the available literature was examined. The initial search terms used in this review were kept broad. An initial search was conducted using the Boolean operators AND, OR and the terms:

"*therapist effect*" OR *"therapist variables"* AND "*IAPT*" OR "*improving access to psychological therapy"*

The search yielded 119 resources in SCOPUS, 5 in MEDLINE, 0 in PsychARTICLES, 4 in

PsychINFO, 39 in PubMedCentral, 1 in CINAHL Plus, 15 in JSTOR, and 2 in Psychology and Behavioural Sciences. In addition to the search using these data bases, academics, authors and senior clinical leaders were contacted either via email, or personally, at conferences. These further searches yielded one unpublished paper (Bruijniks, in review). This paper had not been published at the time of writing (December 2019). A further search of secondary sources, including a snowballing search of all articles, texts and meta-analyses was conducted, yielding no further articles. A total of 186 articles were found, after removing duplicates a total of 112 remained. The selection of articles is shown in the PRISMA diagram in figure 3.1.

Figure 3.1 PRISMA Diagram showing article selection process



3.2.3 The selection of articles

After reading study titles and abstracts, 55 articles were discarded as they did not relate to IAPT, 18 did not relate to CBT, 12 contained no statistical analysis, 8 related to step 2 CBT only, 5 related to children, 1 paper was incomplete (Bruijniks, unpublished) and 1 was not written in English language. This yielded 12 articles for full review (see table 3.2). These 12

articles comprised of 6 naturalistic observational cohort studies and 6 Therapist Effects studies. Each article was reviewed and scored in relation to the literature review question and inclusion criteria. Of the 12 articles, 7 did not meet the inclusion criteria. Waller and Turner (2016) provide a highly useful discussion relating to the therapist variables that may relate to clinical outcome in the delivery of cognitive behavioural therapy. However, their discussion does not specifically relate to IAPT and they fail to include any substantive data analysis or findings. This article was excluded on the basis that it lacked scientific rigor. Clark, Canvin, Green, Layard, Pilling and Janecka (2017) report on the "*transparency*" of data in IAPT in a naturalistic observational cohort study. Whilst they suggest that their sample is in excess of 500,000 patients, they fail to give further details of the sample of interest. Furthermore, they give no details of the therapists that treated the patients. The main recommendation of the article is that IAPT should be adopted in other countries. Close inspection of the article reveals that, unlike its title, there is actually a lack of transparency in the data. IAPT collect large volumes of data about therapists, patients and services (Layard and Clark, 2014) and yet the authors fail to discuss this. They focus on data in a very broad sense, looking at aggregated data reported by groups of IAPT services situated in geographical areas (National Health Service Clinical Commissioning Groups). They present the data using a largely positive stance reporting that recovery rates are in line with those reported by clinical trials and that IAPT is able to capture 98% of outcome data for all patients. Whilst there is a brief discussion about variance in outcomes the authors ultimately present the achievements of IAPT. It could be argued that the authors' primary aim is to market the IAPT concept to other countries in the world. The third study in the review (Branson, Shaffran and Myles, 2015) also failed to meet the inclusion criteria for this review in that the authors discuss the relationship between competence and outcomes in High Intensity CBT trainees. Additionally, their study has significant limitations in that the n=43 trainees who participated in the study were allowed to self-select three therapy sessions to be assessed for competence, using the CTS-R. Self-selection, particularly in the context of a training programme, can be problematic in that the trainees are likely to submit recordings

which they believe to be better examples of their clinical work (Walfish, McAlister, O'Donnell and Lambert, 2012). The authors findings are perhaps unsurprising in that they report that trainees CTS-R scores improved over time, however there was no evidence (p= 0.08) that CTS-R scores correlated with clinical outcome in this cohort of trainees. It is important to note that the authors reported this finding as "limited evidence" rather than no evidence that CTS-R scores related to clinical outcome. Furthermore, the results from this study may not be generalizable to the population of qualified therapists working in IAPT. The fourth article which was excluded from the review was Saxon and Barkham's (2012) therapist effects study. Whilst the authors add to the literature in this well-powered study, close inspection reveals that they repurpose an old data set derived from the Clinical Outcomes in Routine Evaluation (CORE) practice-based evidence national data base, which collected data from patients treated between the years of 1999 and 2008 (Saxon and Barkham, 2012). Therefore, whilst the paper appears to relate to IAPT the data was actually collected prior to IAPT. Two other articles were excluded on the same basis. Firth, Saxon, Stiles and Barkham (2019) and Saxon, Barkham, Foster and Parry, (2016) also use the CORE (1999-2008) database. Pybis, Saxon, Hill and Barkham, (2017) similarly repurpose secondary data in their study. The authors conduct a secondary analysis from data collected as part of the 2nd National Audit of Psychological Therapies. Again, this is a very large data set (n=11750 IAPT patients) however the data set did not contain details of therapists and this article was excluded on that basis. The remaining five articles were critically reviewed using the McMaster Critical Review Guidelines (Law, Stewart, Pollack, Letts, Bosch, Westmorland, 1998). The McMaster guidelines (for quantitative methods) are commonly used by health care professionals when critically appraising studies for inclusion in systematic literature reviews (Bettany-Saltikov, 2012). Table 3.3 shows the McMaster review form for the five articles. The form shows how each article was appraised in relation to whether: the study purpose was clearly stated and relevant, the article included a relevant literature review and an identified gap in knowledge, the research design is clearly outlined, the sample is described and the sample size is stated and justified, outcomes are clearly defined and

measures have sound psychometric properties, the intervention is described and replicable, the findings are reported including the statistical analyses used, the statistical significance is reported and the relevance and impact on clinical practice are discussed (Law et al., 1998). Table 3.3 shows that all 5 articles met the significant majority of the criteria and the Liness et al., (2018) article met all the criteria. Given that the search revealed so few relevant papers, all 5 papers were retained.

Table 3.2 Articles yielded from the review

	Author & year	Type of study	Focus of study	Relevance to review
1	Waller and Turner (2016)	Observational	Therapist adherence	No -excluded
2	Clark et al., (2017)	Naturalistic observational cohort study	IAPT data	No- excluded
3	Branson, Shafran & Myles (2015)	Observational cohort study	IAPT trainees	No -excluded
4	Saxon & Barkham (2012)	Therapist effects	Investigate therapist effects in primary care services	No -excluded
5	Firth et al., (2019)	Therapist effects	Clinic & therapist effects	No - excluded
6	Saxon et al., (2016)	Therapist effects	Therapist effects and patient drop out/deterioration	No -excluded
7	Pybis et al., (2017)	Therapist effects	Secondary analysis of outcome data in IAPT & other psychological therapy services	No -excluded
8	Pereira et al., (2017)	Therapist effects	Therapist effects & therapist resilience/mindfulness/job satisfaction in relation to clinical outcome	Yes- included
9	Delgadillo, Saxon and Barkham, (2017)	Therapist effects	Therapist effects & therapist burnout in relation to outcome	Yes- included
10	Gyani et al., (2013)	Observational prospective cohort study	Variables that predict outcome	Yes - included
11	Liness et al., (2019)	Observational longitudinal study	Relationship between competence & outcome	Yes - included
12	Saxon, Firth and Barkham (2017)	Therapist effects	Relationship between therapist effects, therapy modality dose, no- attendance and outcome	Yes- included

Table 3.3 McMaster Critical Review Form (Law et al., 1998). The table shows the components (left hand column) that were used to critically review each paper. A mark of 'x' denotes that the article met the component. Where there is no 'x' the article did not meet the criteria. A mark of 'NA' shows that this component was not applicable to the article.

	Delgadillo, Saxon and Barkham 2017	**Pereira, Barkham, Kellett and Saxon (2017)**	**Gyani, Shafran, Layard and Clark (2013)**	**Saxon, Firth and Barkham (2016)**	**Liness, Beale, Lea, Byrne, Hirsch and Clark (2018)**
Study Purpose	x	x	x	x	x
Literature	x	x	x	x	x
Appropriate Design	x	x		x	x
Sample Description	x	x		x	x
Sample Size	x	x	x	x	x
Outcome measures validity	x	x	x	x	x
Outcome measures reliability	x	x	x	x	x
Intervention described	NA	NA	x	NA	x
Statistical significance reported	x	x	x	x	x
Analysis method appropriate	x	x	x	x	x
Clinical importance discussed	x	x	x	x	x
Conclusion and implications discussed		x	x		x

3.2.4 Discussion of findings

This literature review sought to understand the findings from published literature in relation to the question 'what therapist variables relate to clinical outcome in High Intensity CBT therapists treating step 3 patients in IAPT'. This search found five articles that met the inclusion criteria. Three of these were therapist effects (see section 1.2.7 for a definition of therapist effects) studies, one was a naturalistic observational cohort study and the other was an observational longitudinal study.

The first substantive attempt to understand the variables that may account for clinical outcome, in IAPT, was undertaken by Gyanni, Shafran, Layard and Clark (2013). This observational, prospective cohort study, analysed the outcome data of 19,395 patients who completed treatment in the first year of the IAPT programme. The patients were treated at step 3 by a High Intensity CBT therapist, a counsellor, or treated at step 2 by a Low Intensity therapist (Psychological Wellbeing Practitioner). Unfortunately, it is unclear how many clinicians were included in the study and patients received either counselling, CBT, or a step 2 intervention. The authors undertake a logistic regression in order to explore three main themes; patient variables, service variables and therapist variables. Their analysis suggested that those patients who score highly on the PHQ-9 and GAD-7 (and therefore the patients with greater severity of symptoms) tended to do less well than patients who had lower scores. In addition, they suggested that patients with a diagnosis of either, generalized anxiety disorder, mixed anxiety and depression disorder or depression were more likely to reach recovery than patients with other diagnoses. Gyani et al., noted that services who provided patients with higher numbers of therapy sessions tended to get better recovery rates. They also reported that services who coincidently had greater numbers of therapist on higher salary bands tended to get better recovery rates. Gyani et al., argued that therapists on higher salary bands would be more experienced and they hypothesized, therefore, that more experienced therapists would get better outcomes. This argument lacks evidence as the authors failed to investigate the data in relation to whether higher salary bands equated

to greater experience. Additionally, it is possible that some services opted to use higher salary bands for all staff, rather than a higher salary band denoting years of experience. It is impossible to rule out that other variables accounted for the higher recovery rates. Furthermore, their assumption that more experienced therapists ger better outcome contradicts the findings from other studies (Beutler, Malik, Alimohamed, Harwood, Talebi, Noble and Wong, 2004 and Webb, DeRubeis and Barber, 2010). Gyani et al., make other weak connections in their discussion relating to therapist variables. They suggest that therapists who deliver interventions that adhere to NICE guidelines (see Chapter 1) are more likely to get better outcomes than therapists who do not. Again, the authors fail to validate that the therapists were delivering NICE approved interventions. They position their argument on the basis that High Intensity CBT Therapists were, on average, achieving better outcomes than counsellors. They report that when treating patients who had a diagnosis of depression CBT therapists achieved a recovery rate of 40 % whilst counsellors achieved 38.3%. A similar difference was reported for patients with a diagnosis of generalized anxiety disorder where patients who were seen by a CBT therapist were 1.324 times more likely to reach recovery. Whilst there is a clear difference between the outcome rates of CBT therapists and counsellors in this study, it is not unlikely that there may have been other variables that account for the difference. These variables may include the types of patients that are allocated to either CBT therapists or counsellors. Given that CBT therapists and counsellors have different backgrounds and training, it is not unrealistic to assume that they may expect to see different types of patients and this may account for the variance in their respective recovery rates. Other variables that might be relevant are differences in training, access to clinical supervision, job satisfaction or rate of pay. Gyani et al., do not comment on these, or any other, potential variables in their assertion that following NICE guidelines predicts higher recovery rates for patients with generalized anxiety disorder or depression. A significant limitation of this article is that whilst the authors draw the conclusion that therapists with higher recovery rates are adhering to NICE guidelines they have not monitored or assessed the therapists work. It might be important to reflect on

the fact that one of the authors, David Clark, has written at length, in a number of texts, about his support of evidence based psychological therapies and NICE guidance. Whilst the authors argument may be a true reflection on the efficacy of evidence based psychological therapies and disorder specific treatment protocols in the NICE guidelines, the authors might have used more robust assessment methods to support their assertions. This is particularly important given that they report the issue of significant variance in outcome between services and therapists.

In article two of this review, Saxon, Firth and Barkham, (2016) comment on the dearth of research investigating why some CBT therapists are more effective than others. The authors highlight research (see section 3.1.) which indicates that some variables such as age, gender, years of experience, adherence to a protocol have either no relationship to outcome or only a very weak relationship. It is noteworthy that, whilst they highlight that little is known about why some CBT therapists achieve better outcomes than others, the authors make no attempt, in this study, to add specificity to what is already known. Saxon, Firth and Barkham use multi-level-modelling (MLM) to analyse the outcome data in order to understand the relationship between therapist effects, therapy dose, therapy modality and patient outcome. Whilst this study was undertaken within IAPT the authors use reliable improvement on the PHQ-9 as the primary outcome measure rather than reliable improvement or recovery using both the PHQ-9 and GAD-7 which are the normal outcome indices in IAPT. It is unknown whether this would have impacted on their estimation of therapist effects and variance in outcome using their multilevel model. Both Saxon and Barkham have used MLM in previous studies (Saxon and Barkham, 2012, Saxon, Barkham, Foster and Parry, 2016 and Firth, Saxon, Stiles and Barkham, 2019) and assert that the results from this particular statistical method produce a model whereby it is possible to see the variance between each therapist and how this impacts on outcome. They use multilevel modelling, in this post hoc analysis of IAPT data, to understand how each therapist differs from each other in terms of the amount of sessions delivered, the therapy modality (CBT or counselling) and the patient outcomes.

Saxon, Firth and Barkham report that the recovery rates in this study ranged from 16 to 76% and that 13% of the therapists achieved significantly better outcomes than the others. These 'better' therapists achieved recovery rates twice that of the 16% of therapists who were deemed less effective. It is unsurprising that in a real-world setting that there will be variance between therapists (Branson and Shafran, 2015, James, Blackburn, Milne and Reichfelt 2001), even when the therapists have undertaken the same clinical training (Waller 2009). Saxon, Firth and Barkham report that overall therapists effect, in their analyses, accounted for 5.8% of patient outcome. In their discussion on session dose the authors report that there was a positive relationship between the amount of sessions and patient outcome, although this was not seen across all therapists. They suggest that the difference between therapists related to the quality of the dose of the therapy being delivered and that, therefore, the best therapists were delivering therapy that was of a higher quality. However, the authors do not expand on what they mean by 'quality of dose' and merely recommend that further research is required to understand this implied phenomenon.

Understanding what Saxon, Firth and Barkham (2017) refer to as the 'quality of the dose' delivered by a therapist is a subject that is explored by Liness, Beale, Lea, Byrne, Hirsch and Clark (2018) in their observational longitudinal cohort study CBT trainee competence. This study explores the relationship between competence and clinical outcome in a cohort of 45 cognitive behavioural therapists both whilst enrolled on an IAPT High Intensity CBT training programme and 12 months post-qualification. The authors use the revised version of the Cognitive Therapy Scale (CTS-R) to assess therapist competence. The trainees are asked to submit a recording of a treatment session three times during their training. The recording is then rated using the CTS-R (see chapters 2 and 4 for further explanation of the CTS-R). They use a cut off of a score of $\geq$ 50% on the CTS-R as a definition of competence 38 therapists scored 50% or above and were defined as competent and 7 scored less than 50% and were deemed less competent. Liness et al., report that whilst the therapists were in training that their CTS-R scores improved over time ($p = < 0.001$). This is an unsurprising

finding given that the main aim of an IAPT training programme is that trainees will be taught how to become more competent. This finding was also reported by Branson, Shafran and Myles (2015) but there was no longitudinal follow up in their study. Liness et al. report that 85% of the cohort remained competent (achieved a score of ≥ 50% on the CTS-R) 12 months post-qualification. There is surprisingly little variance in this cohort, and this may be due to two factors. The first factor is the research setting; this study was conducted at the Institute of Psychiatry, Kings College, London. Arguably this is one the better IAPT training programmes in that it has a higher intake of Clinical Psychologists and the training programme is highly sought-after, making gaining a place highly competitive (Liness et al., 2018). This may enable this particular IAPT training programme to select higher caliber candidates, thus reducing the variance in outcomes. The second factor that may reduce variance in outcome is that this study permitted therapists to self-select therapy recordings to submit for assessment. The problem associated with self-selection of therapy recordings is that the process is likely to lead to self-selection bias. That is to say, therapists may select their best therapy sessions (Barber, Shapless, Klostermann and McCarthy 2007). Whilst the self-selection of recordings may be problematic Liness et al. employ robust inter-rater reliability training and report a high ICC of 0.80 between raters. Secondly, 18% (n=22) of the recordings were second marked by an external rater who was blind to the status of the therapist. Analyses of covariance showed no significant relationship ($p = < 0.06$) between CTS-R score and clinical outcomes whilst the therapists were in training. However, 12 months post-qualification there was a significant relationship between CTS-R score and outcome ($p = < 0.05$). It is important to note that the authors used reliable improvement on the GAD-7 (i.e. a reduction of 4 points between first and last score) as the primary outcome measure in their analysis. It is unclear why this is the case and it would have been helpful to know what the strength of the relationship would have been, between competence and outcome, had the authors had used the PHQ-9 or the IAPT definition of recovery as the primary outcome measure. Furthermore, the authors report that there was no difference in

the therapists' outcomes whilst they were training and 12 months after they had completed training. This latter finding is surprising in that it might be expected that the therapists' outcomes would improve after they had qualified. However, the recovery rates reported in this study were only based on 16 patients (8 whilst they were training and 8 post-qualification) so this may be an issue that relates to a small sample size. The authors conclude that they had insufficient evidence that there is a relationship between competence and outcome but highlight that this may be due to the lack of variance in the cohort. Clearly this article does much to celebrate the quality of training at the Institute of Psychiatry but, as a result, the findings may be less generalizable.

The two remaining articles in the findings from this systematic review both use multilevel modelling to estimate therapist effects in IAPT. Pereira, Barkham and Saxon (2017) undertake a feasibility study on the relationship between therapist resilience, mindfulness and clinical outcome. Pereira (as part of her PhD) works with Barkham and Saxon using, MLM to understand the relationship between the two independent variables (therapist resilience and mindfulness) and clinical outcome, using reliable improvement on the PHQ-9 as the primary outcome measure. Their study looks at two new variables (therapist resilience and therapist mindfulness), which very much add to the literature. The authors use two self-report questionnaires to assess both variables. They use the results from these, together with patient outcomes in their MLM. Periera, Barkham and Saxon report a significant variance between therapists' outcomes and differentiate between the most effective and the least effective therapists. They estimate that therapist effects account for 6.7% of variance in patient outcome and they present a strong argument that high levels of therapist resilience and mindfulness are important factors in the most effective therapists ($p = < 0.005$). Out of the 42 therapists in this cohort, 11 are Psychological Wellbeing Practitioners (PWP), 19 are counsellors and only 12 are High Intensity CBT therapists. The authors report that both CBT therapists and counsellors have roughly equivalent clinical outcomes and both groups tend to be more mindful than PWPs. They suggest that this is

because PWPs are not exposed to mindfulness in their training. This argument is somewhat flawed because mindfulness is not on the IAPT training curriculum for High Intensity trainees either. In fact, the measure the authors used (Mindfulness Attention Awareness Scale, Brown and Ryan, 2003) does not assume prior training in mindfulness and is a commonly used self-report measure used to assess traits in patients (Brown and Ryan, 2004). Additionally, the use of self-report measures alone rather than including a clinical assessment might be questioned. However, it is likely that undertaking a clinical assessment may have deterred participants and may be ethically and logistically challenging. Despite the challenges, the authors new and innovative findings on what variables may account for the variance in clinical outcome add a new theme to the literature. If therapist resilience and mindfulness are a factor that leads to better patient outcome then this has implications for both trainee selection, therapist training, continuing professional development and patient outcomes.

In the final article of this review the authors also look at therapists' emotional factors as a predictor variable of clinical outcome. In this study Delgadillo, Saxon and Barkham look at the associations between therapists' occupational burnout and the clinical outcomes of their patients. Despite the rising incidence of occupational burnout amongst therapists in IAPT (Westwood, Morison, Allt, Holmes, 2017) there is very little literature investigating the impact of burnout on clinical outcomes. As in the previous study, the authors used a self-report measure (Oldenburg Burnout Inventory, Demerouti, Bakker, Nachreiner and Schaufeli, 2001) to rate occupational burnout. Perhaps unsurprisingly, the authors found a significant ($p = < 0.05$) relationship between occupational burnout and clinical outcome with an overall therapist effect of 5%. Of the n = 49 therapists included in this study, only 21 were High Intensity CBT therapists. Unfortunately, the authors do not provide a specific breakdown of their findings, so it was unclear how the finding related specifically to High Intensity CBT therapists. Additionally, it is unclear how much of the overall therapist effect was due to occupational burnout in the whole cohort. Regardless of these issues this article examines a

previously unexplored variable and the implications of the findings, in relation to providing adequate support to therapists in order to improve clinical outcomes, are not insignificant. It is these implications that drive researchers to learn more about how therapist variables impact on service delivery and recovery rates. It is clear from the research on general therapist effects that the therapist may be accountable for up to 8.6% of the variance in recovery rates (Crits-Christoph, 1991). However, there remains a great deal of conjecture about whether this figure has been overestimated (Brown et al., 2005, Wampold and Brown, 2005). Alternatively, it is possible that previous studies are underestimating the significance of therapist variables in relation to patient outcome (Baldwin and Imel, 2013). This may occur because most studies are not set up to specifically study therapist effects. However, the three therapist effects studies included this review have all been designed as such.

3.2.5 Sample size

There has been some debate about the importance of sample sizes in the study of therapist variables and therapist effects (Schiefele, Lutz, Barkham, Rubel, Bohnke, Delgadillo, Lambert et al., 2017). The studies in this review were included because they met the recommendations for sample size in these types of studies. Figures 3.2 and 3.3 show the sample size (therapist and patient) for each study. However, given that in each of these studies the sample of interest is the therapist, it might be argued that that whilst the studies were adequately powered in relation to the numbers of patients that were treated but that the sample size of therapists might have been larger.

Figure 3.2 Sample size – therapist (Gyani et al., fail to disclose therapist n)

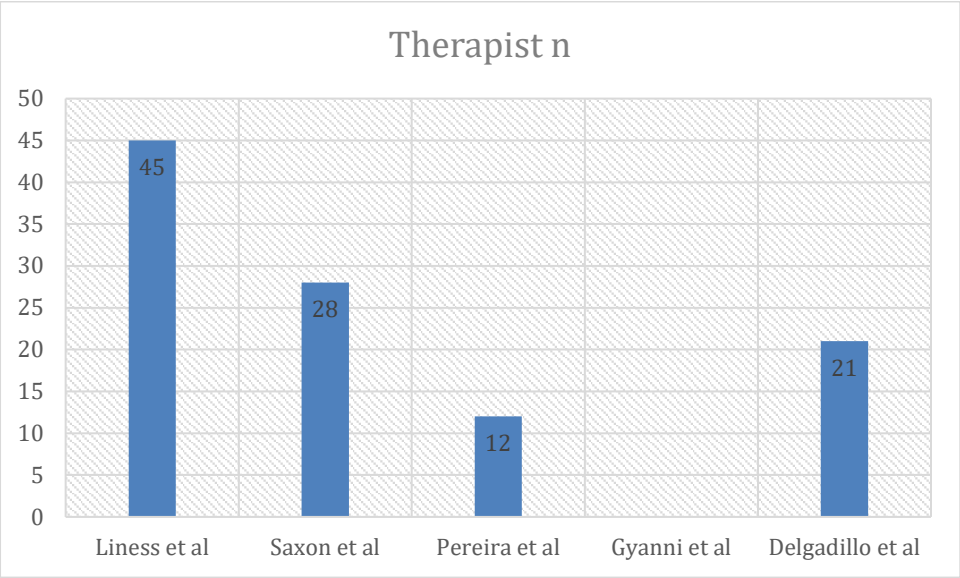
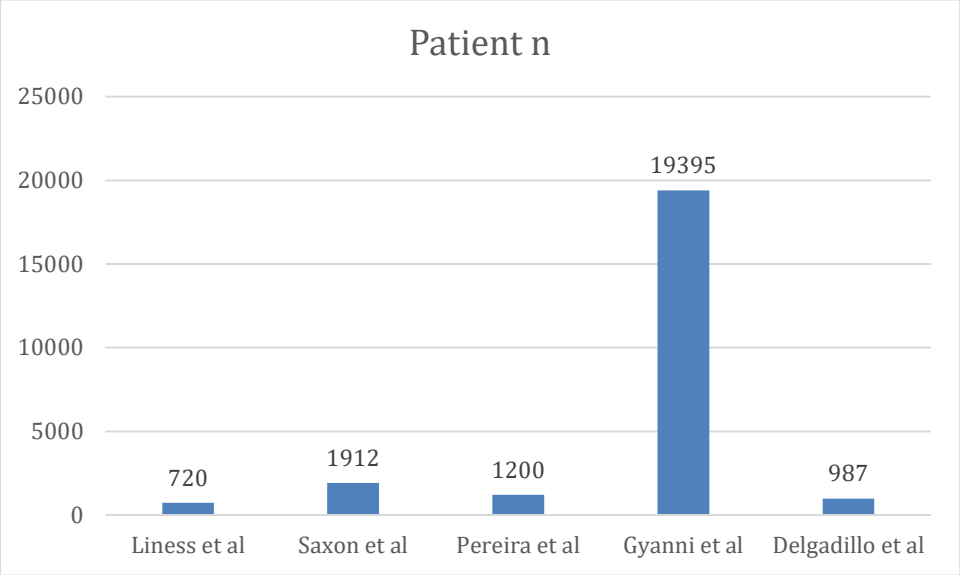


Figure 3.3 Sample size – patient



3.2.6 Limitations of the systematic literature review

One significant limitation of this systematic literature review is that it was conducted alone. Therefore, in the absence of any collaborators it is possible that the researcher may have been subject to her own bias in relation to whether or not articles should have been included or excluded in this review. Furthermore, the absence of collaborators may impact on the methodological processes incorporated into the systematic literature review. Whilst an a priori protocol was developed for this review the absence of collaborators means that human error cannot be excluded.

Additionally, risk of publication bias cannot be eliminated. Whilst significant attempts were made, over a period of three years, to contact authors and key opinion leaders to investigate whether there were any unpublished articles, this yielded just one paper that had not been published. It remains unclear whether other articles exist that might have otherwise been included in this systematic literature review. Furthermore, whilst two of the studies included in this review had modest findings and were less likely to be subject to publications bias, two papers are significantly aligned with the IAPT agenda and it might be argued that reviewers and editors would look more favourably on these.

3.2.7 Summary

It is clear from the articles reviewed that therapist variables are having some impact on patient recovery. However, there is a lack of agreement about to what extent this is and, more importantly, what it is that therapists are doing that may be impacting on their patients. Both therapist effects studies and naturalistic observational studies offer a great deal in helping answer these questions. However, there are some common limitations within most, if not all, of the studies conducted to date. These include small sample sizes of High Intensity CBT therapists, lack of access to live therapy recordings and allowing therapists to self-select therapy recordings. Future research would be needed to address these limitations.

3.3 RECOMMENDATIONS FOR FURTHER RESEARCH

The preliminary scoping review and systematic review of the literature have revealed considerable agreement in relation to the recommendations for future research. These include increasing the sample size of therapists, heterogeneity of therapy models and therapist minimum training and improving access to live therapy material (Baldwin and Imel, 2013). These issues, together with recommendations for further research, are illustrated in table 3.4. It is evident that whilst it is widely accepted that variance between therapists is partly responsible for variance in outcomes what is less clear is why this is. The literature suggests that no further research is required in relation to therapist demographics or the therapist's ability to develop and maintain a therapist alliance as it is agreed that therapist demographics are not related to outcome and the therapeutic alliance is related to outcome. What is less clear is to what extent therapist competence and adherence are directly related to outcome. Each of the studies discussed highlight the issue of closely scrutinizing the way therapists work in order to begin to study this question in more depth. Section 3.4, below, discusses how the findings from this literature review have been used to inform the research questions addressed in this thesis.

Table 3.4 Recommendations for further research in current articles

Article	Limitations	Suggestions for further research
Ehlers et al., (2013)	Small sample size	Add measures of therapist competence and adherence Increase sample size
Saxon, Firth and Barkham (2017)	PHQ-9 is the only primary outcome measure and no diagnosis is recorded so it is unclear if patients with a diagnosis other than depression benefited. No measure of alliance or adherence	Data set should include: Outcome measures, patient diagnosis, therapist factors, therapist characteristics
Branson, Shafran and Myles (2015)	Missing outcome data. Self-selection of therapy recordings. Only based on trainees who are likely to improve over the duration of their training.	Do not allow therapist to self-select therapy recordings Use more raters to assess competence (CTS-R)
Laska, Smith, Wislocki, Minami and Wampold (2013)	No adherence monitoring. Unable to assess what the better therapists were doing in their sessions. No measure of alliance. Limited access to patient outcome data Small sample size and mix of trainees and qualified clinicians	Ensure therapists are trained in the same therapeutic model. Use measures of adherence Have access to therapy recordings and all outcome data
Ginzburg, Bohn, Hofling, Weck, Clark and Stangier (2012)	Selected therapists who already showed high levels of adherence and competence Only used one recording to assess adherence	Use real-world clinical setting. Assess adherence using more than one therapy recording Increase sample size
Laska, Smith, Wislocki, Minami and Wampold (2013)	No adherence monitoring. No measure of alliance. Used supervisors to rate competence. No access to therapy recordings	Measure adherence Have access to therapy recordings
Barnfield and Beaumont (2007)	No outcome data	Use outcome data. Use multiple measures of compliance. Conduct inter-rater reliability analyses on raters
Weck et al., (2016)	Small sample size	More research in real-world settings Consider whether adherence is more important in treating anxiety and competence is more important when treating depression

3.4 RESEARCH QUESTIONS FOR THIS THESIS

Following the findings from this literature review, it is clear that a small but significant body of work exists in relation to therapist variables and their relationship with outcome in IAPT. The body of work in therapist effects present some agreement that between 3 and 8% of variance of outcomes in IAPT can be accounted for by therapist effects. What is less clear are the therapist variables that relate to therapist effects. That is to say, what therapist variables are related to outcome in High Intensity Therapists treating patients at step 3 in IAPT? There remains a significant gap in the literature in this area. This is most likely to be due to the lack of availability of recordings of therapy sessions.

This research aims to provide an original contribution to knowledge, building on the presented body of research and utilizing the recommendations for future research, by asking the following research questions:

1. Does therapist age, years of experience or their core profession correlate with clinical outcome in a step 3 (High Intensity) IAPT service?

   Null Hypothesis: Therapist age, years of experience and therapist core professions have no relationship with clinical outcomes in a step 3 (High Intensity) IAPT service

   Alternative Hypothesis: Therapist age, years of experience and therapist core profession are directly related to clinical outcomes in a step 3 (High Intensity) IAPT service.

2. Do therapists who have completed the IAPT training programme achieve better clinical outcomes than those who have not?

   Null Hypothesis: There are no differences between clinical outcomes between therapists who have completed an IAPT training programme and those that have not.

Alternative Hypothesis: Therapists who have completed an IAPT training programme will achieve higher clinical outcomes than therapists who have not completed an IAPT training programme.

3. How much does therapist competence explain the variance in patient outcomes in a step 3 (High Intensity) IAPT service?

   Null Hypothesis: Therapist competence does not explain any variance in outcomes a step 3 (High Intensity) IAPT service.

   Alternative Hypothesis: Therapist competence is related to clinical outcomes a step 3 (High Intensity) IAPT service.

4. How much does therapist adherence to evidence-based protocols explain the variance in outcomes in a step 3 (High Intensity) IAPT service?

   Null Hypothesis: Therapist adherence to an evidence-based protocol does not explain any variance in clinical outcomes a step 3 (High Intensity) IAPT service.

   Alternative Hypothesis: Therapist adherence to an evidence- based treatment protocol is related to clinical outcomes a step 3 (High Intensity) IAPT service.

3.5 CONCLUSION

This chapter began with a preliminary scoping review of the literature in relation to the therapist variables that may relate to outcome more broadly in psychological therapy. The themes that emerged from the scoping review were then discussed, including therapist demographics, therapist competence, therapist adherence, therapist ability to develop and maintain a therapeutic relationship and therapist effects. This review of the wider literature

identified that whilst there is some conjecture about what therapist variables relate to outcome, the issue remains unclear because the studies reviewed have small sample sizes and/or researchers have had little or no access to live therapy recordings. This theme also emerged in the systematic literature review which examined therapist variables and their relationship with outcome in High Intensity CBT therapists treating patients at step 3 in IAPT. Whilst there were a number of articles that explore the variance in outcomes therapists, many of these were therapist effects studies. These robust and well-powered studies add a great deal to what is known about the variance in outcomes in IAPT, but a significant gap remains in relation to what therapist variables are related to outcome within the overall therapist effect. This gap in knowledge is due to the lack of access to recordings or transcripts of therapy sessions and this has limited new learning in relation to the variance in recovery rates in IAPT. Many of the studies reviewed have not had access to therapy recordings at all and, of those that did, the availability of therapy recordings was limited and were usually self-selected by the therapist. This chapter draws from the recommendations made in the articles reviewed suggesting that future research uses sufficient recordings or transcripts of therapy sessions. Internet Enabled CBT (IECBT), as described in Chapter 2 of this thesis, results in a transcript for every therapy session delivered. Therefore, the delivery of High Intensity CBT, online using written communication, (IECBT) provides a new method of understanding what CBT therapists are doing with their patients and how that is related to clinical outcomes. This method of delivering CBT also enables transcripts to be randomly selected for review.

This chapter concluded with the research questions for this thesis. Chapter 4 will outline the methodology, research design and data analysis for this research.

CHAPTER FOUR: METHODOLOGY

“there is nothing so practical as a good theory” (Lewin 1951 p.169)

Chapter 3 of this thesis discussed a review of the literature relating to how therapist variables relate to outcome in the wider psychological therapy literature and in the IAPT literature. The chapter highlighted a significant gap in the knowledge relating to the therapist variables that are associated with clinical outcome in IAPT. It was argued that the gap in the literature exists because in order to address this research question, it is necessary to have access to sufficient numbers of recordings or transcripts of therapy sessions. Chapter 3 argued that High Intensity (step 3) CBT delivered in IAPT via Internet Enabled CBT (IECBT) provides the opportunity to randomly select transcripts in order to learn what therapists are doing with their patients that might relate to outcome. Chapter 3 concluded with the research questions for this thesis.

This chapter presents an analysis of the methodological issues relevant to this study. It includes a rationale for the approach and a review of the theoretical frameworks traditionally used within contemporary cognitive behavioural therapy practice. This chapter goes on describe the research design and statistical analyses that will be used in this research.

4.1 THE PHILOSOPHY OF EVIDENCE-BASED PSYCHOLOGICAL THERAPY

Evidence-based psychological interventions arose in the United Kingdom (Layard and Clark 2009). The underpinning premise was that clinical outcomes would only improve with the acquisition of and utilisation of contemporary learning (Eysenck, 1966). Epistemic terms such as ‘evidence’ ‘grounds’ and ‘warrant’ have become synonymous with the CBT literature. There is some debate relating to how, and when, researchers can claim they know that a theory is likely to be plausible and to what extent is it possible to say that one theory is

superior to another (Williams, 2015, Mollon, 2010, Layard and Clark, 2014). These epistemic arguments most frequently relate to research methodology. There has been some criticism about a number of CBT studies (Andrews, 2000, Wampold, Goodheart and Levant, 2007, Lowenthal and Proctor,2018, Wampold and Imel, 2015 . The primary argument that is presented is that methodologically most are weak (Wampold Fluckiger, Del Re, Yulish, Frost, Pace, et al., 2016) and, as a result, the studies make inflated claims. For example, one author suggests that claims are not “supported by the evidence that is needed” (Wampold et al., 2016, p.29). Again, the epistemic term ‘evidence’ is used to describe how researchers are required to present assumptions, data or knowledge that corroborates their claims. This focus on evidence in psychological therapy seems to have originated from Eysenck’s studies on the efficacy of behaviour therapy in the 1950s and 1960s (Eysenck 1952, 1961, 1966). A counterargument that aims to address the critics of positivist research in CBT support the biomedical model of science in health care that advocates that researchers should use robust scientific methods to investigate the efficacy of psychological therapies and that interventions that lack this type of evidence should not be adopted in clinical practice (Layard and Clark, 2014). The large number of CBT efficacy studies and subsequent meta-analyses (Smith and Glass,1977, Smith Glass and Miller, 1980, Hunt 1997, Mann, 1994), using a positivist paradigm, belies the call to arms to use an alternative approach. The resulting studies in Europe, Australia and the United States of America have tended to synthesise the evidence and make their own claims about which aspects of CBT are most appropriate for specific patient populations. Consequently, evidence-based psychological therapies, especially cognitive behavioural therapy, have been more widely adopted, particularly in the United States of America and the United Kingdom (McHugh and Barlow, 2012). The term CBT was first used in the literature in the mid 1970s with the first efficacy trials published at the end of the 1970s (Beck et al., 1979). The focus on building empirical evidence, from randomized controlled trials, and a quest for knowledge about what works for whom, has led to the large evidence base for CBT (Layard and Clark, 2014). The dominance of a positivist stance has placed CBT in alignment with a biomedical approach to

research making it more likely that it would be recommended by establishments such as the National Institute of Health and Care Excellence (NICE). CBT has followed a very similar research methodology to medical science whereby interventions are first tested in research settings before being implanted in clinical practice (Layard and Clark, 2014). This positivist stance has become entwined in the practice of CBT whereby therapists are encouraged to be scientist practitioners (Grant and Townend, 2007) reflective practitioners (Bennett-Levy, 2006) and an advocate of practice-based evidence (Westbrook and Kirk, 2005). Therapists are encouraged to learn from every patient they treat, exploring the mechanisms of change and developing hypotheses about why a patient may or may not have recovered (Westbrook and Kirk 2005, Persons, Bostrum and Bertagnolli, 1999 and Fairburn et al., 2009). Given the dominance of positivism in the historical development of CBT in its research, the training of therapists, and in clinical practice itself, it is perhaps unsurprising contemporary CBT research continues to use positivist methodologies, such as randomised controlled trials, pre- and post-intervention studies, correlational studies and related meta-analyses. Arguably, CBT has been 'rewarded' for its positivist stance through the recommendation of bodies such as NICE and the consequential widespread adoption of the model (Williams, 2015, Chambless and Ollendick, 2001). This phenomenon is the basis for some of the criticisms of CBT in that positivism and the use of numerical data, in the form of outcome measures and recovery metrics, might not be the only source of knowledge. There is a growing body of literature that argues that an interpretive paradigm might offer CBT the opportunity to learn more about service-based cultures and patient/ therapist experience (Williams, 2015, Mollon, 2009, 2010). Whilst not to suggest that this argument is incorrect, as it likely that both paradigms have something to offer, it should be noted that this literature tends to be written by those who are not orientated to CBT and come from other traditions such as psychodynamic psychotherapy. Despite the counterarguments, the focus on positivist empiricism continues to be central to the delivery of CBT (Clark, 2014). The term 'collaborative empiricism' was used by Beck, Rush, Emery and Shaw (1979) to describe the clinical practice used by a CBT therapist as they enable the patient to learn experientially.

Therapists are taught to rely on scientific methods to guide their treatments with patients and are encouraged to reflect on why an intervention has been either effective or ineffective (Bennett-Levy and Finlay-Jones, 2018). The premise of establishing and subsequently testing a hypothesis runs in parallel with developing more effective treatments (as in efficacy research) and at the point of delivering treatment to a patient. It is therefore understandable, that within this tradition, cognitive behavioural therapists tend to assert that they are scientist practitioners and that most CBT research sits within a positivist paradigm. The primary aim is to learn, what works, for whom and why (McHugh and Barlow, 2012). CBT has always been based on the epistemological concept of empiricism (Clark, 2014), whereby learning is believed to be derived from experience. This concept has been termed the 'outside in' philosophy (Gipps, 2012) which underpins all cognitive behavioural theory.

4.1.1 Cognitive behavioural theoretical frameworks

Several theoretical models and frameworks have emerged within CBT. These tend to be ways of understanding or characterizing phenomena such as diagnoses, cognitions, emotions, physiology and behaviour (Gelder, 2012). The two main theoretical frameworks used in CBT are: Beck's Cognitive Theory (Beck,1967) and Watson's Behavioural Theory (Watson, 1913). Both these models relate back to the 'outside in' philosophy in that they stipulate that human learning is derived from external experiences that are then internalised. For example, Beck's Cognitive Theory (figure 4.1) is based on how external antecedents (experiences) are cognitively appraised leading to the formation of beliefs about oneself, the world and others (Beck's Cognitive Triad, see figure 4.2).

Figure 4.1 Cognitive Theory (Beck, 1967)

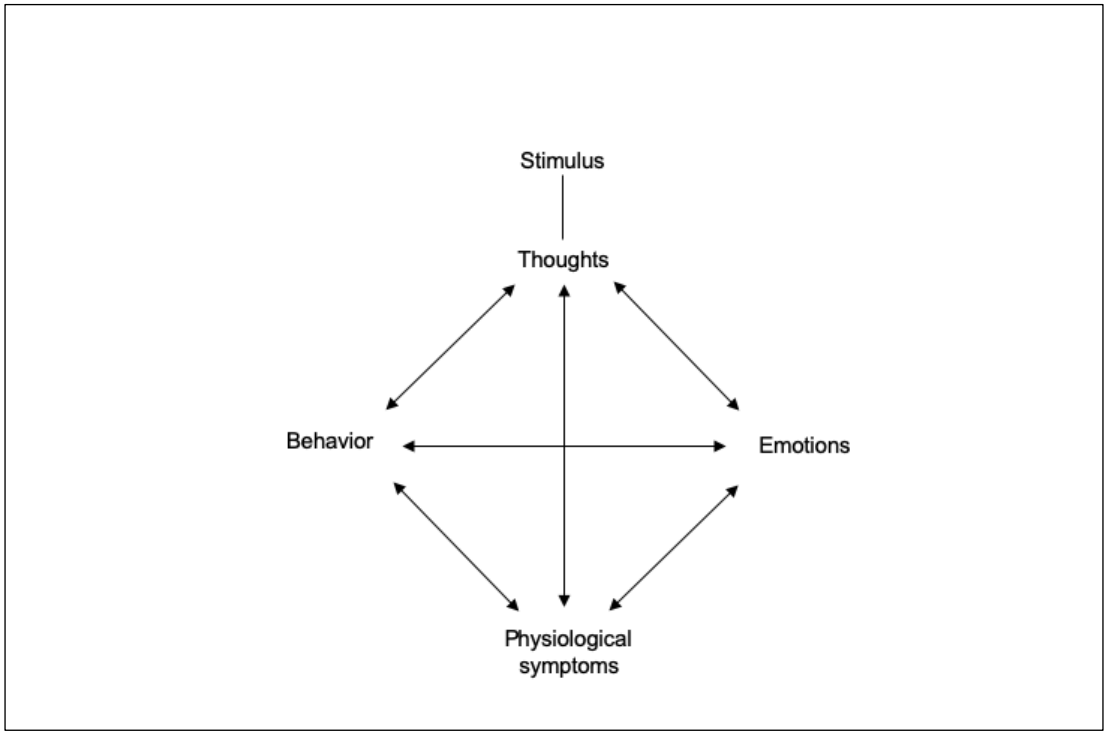


Figure 4.2 Beck's (1979) Cognitive Triad

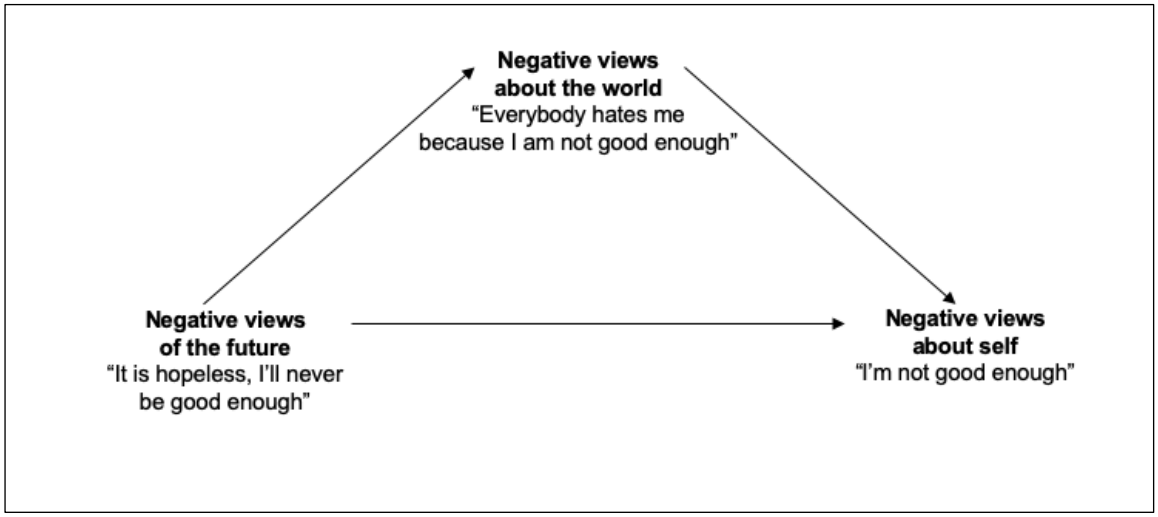
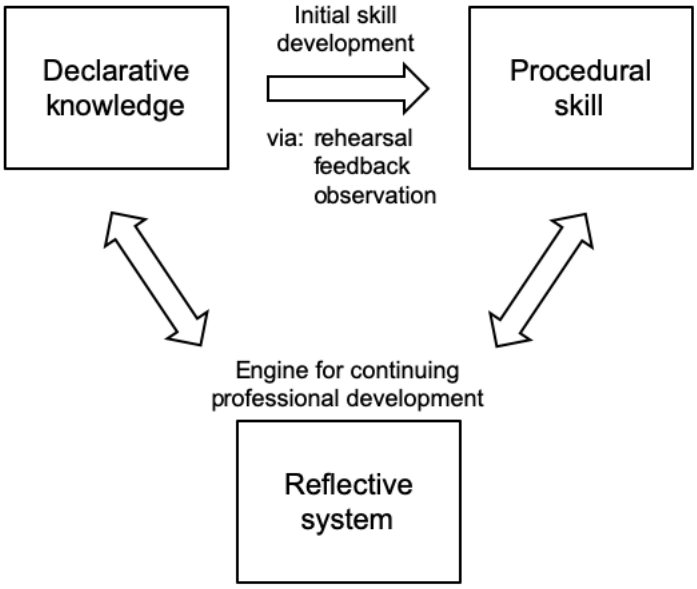


Figure 4.3 The Cognitive Model of Skill Acquisition (Bennett-Levy, 2006)



These theoretical models have informed conceptual frameworks that have been embedded into IAPT training programmes. Arguably the most commonly used framework is Bennett-Levy's (2006) cognitive model of skill acquisition. This model (see figure 4.3) describes the iterative processes that are involved in skill acquisition and the development of clinical expertise. Bennett-Levy places a significant emphasis on the sequential and cyclical process involved in the development of clinical expertise. This process starts with the acquisition of knowledge followed by skill practice, feedback, self-reflection and then returning to start the process again as a cycle of continuous professional development. A more advanced iteration of this model has been described in Chapter 2, section 2.10 of this thesis. Bennett-Levy's conceptual model underpins the processes implicit within the research described in this thesis and is congruent with the Scientist Practitioner approach discussed in section 4.1 of this chapter.

4.1.2 Rationale for the approach

Whilst it might be argued that alternative methodologies could offer much in terms of providing rich data and a variety of perspectives (Thorpe, 2002) such approaches would make it difficult to understand the relationship between therapist variables and clinical outcome in the context of an audience who may be highly critical of anything other than the presentation of findings resulting from quantitative statistical methods set within a positivist framework. In addition, an alternative paradigm might have failed to mediate for the subjective opinions that undoubtedly occur in the context of insider research (Simmons, 2007) which might have detrimentally impacted on the results of this research. A challenge for insider research, conducted by a clinician in their field of expertise is to mediate for researcher bias and the impact this may have on the research results. A positivist paradigm was selected so that an attempt could be made to reduce the impact of the researcher's subjective opinion and prior knowledge. The intention was to take all reasonable steps to maintain objectivity in order to scientifically examine the relationships between therapist variables and patient outcomes. The design attempts to use as much scientific rigour as was reasonably possible, in the context of a doctoral research study in order to increase the confidence that the findings from this research were derived from a scientific approach. The purpose of this research was to explore bivariate relationships and multiple relationships and predictions among variables. This study used bivariate correlations (r) to assess the relationships between all pairs of variables in the study. A multiple correlation (R) was used to assess the relationship of; key therapist demographics (age, gender, core profession etc), therapist ability to deliver CBT with fidelity to the model and therapist ability to adhere to an evidence protocol with clinical outcome (recovery). This research used a multiple linear regression model to establish the variables that predict clinical outcome (Y) from the therapist's ability to a) deliver CBT and b) adherence to an evidence-based protocol (X). In addition, this research sought to understand the extent, to which, two independent variables (fidelity to the CBT model and adherence to an evidence-based protocol) individually and collectively predict clinical recovery. Each independent variable was assessed for its

significant prediction of clinical outcome, and the combination of both independent variables will be evaluated as to how they predict clinical outcome.

4.2 RESEARCH DESIGN

This research used a naturalistic observational study design to understand the relationship between therapist age, gender, experience, training, core profession, competence and ability to adhere to an evidence-based protocol and clinical outcomes. This research employed both bivariate and multivariate analysis to explore the relationship between these therapist variables and clinical outcome.

4.2.1 Setting

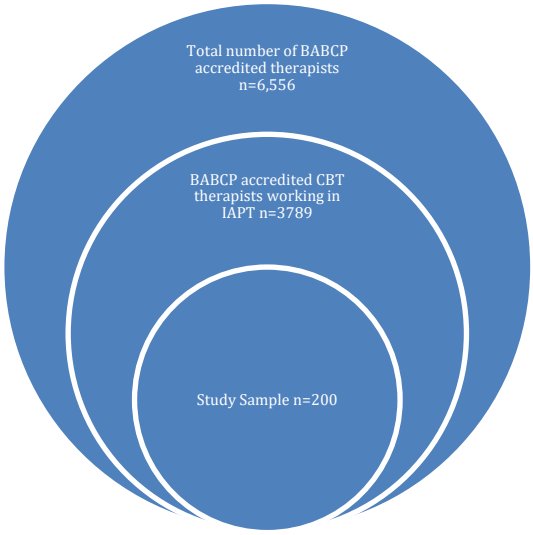
This research was conducted in an Improving Access to Psychological Therapy Service provided by Ieso Digital Health (see www.iesohealth.com). The service is commissioned by the National Health Service (NHS) to provided cognitive behaviour therapy online, using synchronous written communication. The researcher is employed by Ieso Digital Health as the Chief Clinical Officer but has no direct line management responsibility for the therapists working within the service. The service treated 12,000 patients within IAPT in 2018, 10,500 of which were treated at step three.

4.2.2 Recruitment of therapists and sampling method

The sample in this study was selected using a homogenous sampling method (Shadish, Cook, & Campbell, 2002), In this sampling method the characteristics of typical Improving Access to Psychological Therapy (IAPT) High Intensity CBT therapists were defined using demographic data, available in the public domain, from both IAPT and the British Association of Behavioural and Cognitive Therapy (BABCP). In order that the findings from the study might be generalisable the researcher ensured that only those therapists who

matched these characteristics were invited to participate in this study. The therapists being studied in this sample are British Association of Behavioural and Cognitive Psychotherapy (BABCP) accredited High Intensity Cognitive Behavioural Therapists who had completed treatment with at least ten patients (see section 4.2.3 for details regarding the minimum number of patients each therapists had treated) between April 2017 and April 2018 at the online CBT service Ieso Digital Health. Using BABCP (the accrediting body for CBT therapists in the UK) accredited therapists ensured that each therapist had met the minimum training standard as established by BABCP. 474, who met the criteria, were approached to participate in this study. Of the N=474, 237 therapists consented to participate. Of the 237, 37 of these therapists had been assessed by a rater that had been verified, post hoc, as consistently giving higher ratings than the other raters (see Chapter 5 for further details about inter-rater reliability). Therefore, n = 200 therapists, who had consented to participate, were included in this research. This represents 5.28% of BABCP accredited therapists working within IAPT in 2018. Figure 4.4 illustrates how the study sample is sited in the population of IAPT therapists.

Figure 4.4 The study sample in comparison to the population of BABCP accredited therapists



4.2.3 Sample Size

Sample size calculations require the availability of historical data from similar studies (Shieh, 2017). Chapter 3 of this thesis outlined that there are no similar studies in this field at the current time and, therefore, guidance was drawn from the literature on therapist effects studies. The median sample size for previous therapist effects studies is n = 9 therapists, with a range of 2 – 581 therapists (Johns, Barkham, Kellett and Saxon, 2019). A small sample size is commonly cited as a limitation in the majority of the therapist effects studies. Schiefele, Lutz, Barkham, Rubel, Bohnke, Delgadillo, Kopta, et al., (2017) provide recommendations for sample size in their table for real-world research studying therapist effects. Whilst this research is not a therapist effects study there are some similarities in that this research uses naturalistic data in a similar setting, using outcome as a dependent variable. Schiefele et al., (2017) offer useful guidance and in the absence of historical data drawn from therapist variables studies this guidance seemed the most robust sample size advice available. Schiefele et al., (2017) suggest that where the sample size is ≤ 40 therapists then each therapist should have treated at least 4 patients, but in order to increase the confidence interval to 95% it is preferable for each therapist to have treated 30 patients. Where more therapists (i.e. more than 100 therapists) are recruited Schiefele et al., (2017) suggest that the number of patients treated per therapist could be as low as 5. Other guidance for correlational studies suggests a rule of thumb that sample sizes between 30 and 500 are adequate except in multivariate research where the sample size should be 10 times as large as the number of independent variables used in the study (Sekaran and Bougie, 2016). As there are 19 independent variables in this study (see table 4.1) a sample of 200 therapists was deemed sufficient for the multivariate analyses. Additionally, based on the guidance from Shiefele et al., the 200 therapists were required to have treated a minimum of 10 patients.

4.2.4 Raters

Six clinicians (including the researcher) had previously rated therapist transcripts. All of the raters had recent experience of teaching and assessing at a post-graduate, IAPT High Intensity CBT training programme. The raters undertook inter-reliability training for assessing therapist competence, using the CTS-R and adherence to evidence-based protocols using the Roth and Pilling (2006) framework (see chapter 2 for a detailed account of the Roth and Pilling framework).

4.2.5 Patients

The Internet Enabled CBT (IECBT) platform is owned and managed by the company Ieso Digital Health (see www.iesohealth.com). The service delivers Cognitive Behavioural Therapy to patients within the United Kingdom's IAPT programme. IAPT services offer evidenced based interventions for patients who present with common mental health problems such as anxiety disorders and depression. Patients may either self-refer themselves to an IAPT service or be referred by their General Practitioner or another primary health care clinician. The service uses a stepped care model whereby patients are triaged by the service and allocated to either a step 2, low intensity, intervention (mild level of severity) or a step 3, high intensity, intervention (moderate to severe presentation). Step 2 interventions are provided by Psychological Wellbeing Practitioners and step 3 interventions are provided by, more highly qualified, High Intensity CBT therapists. This research focuses only on cognitive behavioural treatments delivered at step three. Step 3 patients are a clearly identifiable population within an IAPT setting and results from this population are regularly compared in analysis of IAPT data (Clark, 2018). The following IAPT inclusion and exclusion criteria were applied for patients treated at step three:

IAPT inclusion criteria

All patients aged 18+ who present with a common mental health disorder such as (but not exclusive to): anxiety, depression, post-traumatic stress disorder, obsessive compulsive disorder, generalised anxiety disorder, agoraphobia, panic disorder, specific phobias, health anxiety, unexplained medical symptoms and long-term medical conditions.

IAPT exclusion criteria

All patients who are or who become actively suicidal or present as a risk to others and patients who are experiencing symptoms of psychosis, hyper-mania, severe, cognitive impairment, severe personality disorder or severe learning disability. In addition to the routine IAPT exclusion criteria Ieso Digital Health operate the following criteria:

- Patients who do not have access to an Internet enabled device or an Internet connection.
- Patients who have a low level of literacy i.e. patients who cannot write or read emails or texts
- Patients who are visually impaired and are unable to write on or read from a computer and do not have access to appropriate assistive technology for the visually impaired.
- Patients who do not speak English.

4.2.6 Treatment provided by the therapists

The therapists provided CBT following the disorder specific protocols used as standard within an IAPT CBT service. All IAPT services are required to deliver evidence-based disorder specific interventions for depression and anxiety disorders. These follow NICE guidance and the Roth and Pilling (2007, 2008) competencies framework, and also form part of the clinical training for all BABCP accredited CBT therapists (further information about NICE guidelines and the Roth and Pilling competencies framework can be found in Chapter

2 of this thesis). Following NICE guidance and the Roth and Pilling competencies framework, individual therapists may select any appropriate treatment protocol when working with a patient. As this is a naturalistic observational study, and the data was analysed post hoc, this research had no influence on which treatment protocol the therapists selected. The number of treatment sessions offered to each patient was at the discretion of the therapist and their Clinical Supervisor and generally depended upon the level of severity of the patient's presenting problem. On average patients treated at step 3, face-to-face, within IAPT complete treatment within four months and have between 8 and 14 treatment sessions. The mean number of sessions in face-to-face is 6.8 (NHS Digital, 2018). The mean number of sessions for patients treated in this study was 7. Patients routinely completed three self-administered questionnaires before each therapy appointment. This is normal practice within IAPT services. Routine measurement is part of the IAPT minimum data set and includes: the GAD-7 (Spitzer et al., 2007), PHQ-9 (Kroenke et al., 2001) and Work and Social Adjustment Scale (Mundt, Marks, Shear and Griest, 2002). In addition, IAPT require patients to complete a patient experience questionnaire (PEQ) at the end of treatment.

4.3 ETHICAL CONSIDERATION

This study was conducted in compliance with a Research Ethics Committee (in this case) favourable opinion. Ethical approval for this trial was granted on 17th January 2018 by the Departmental Research Ethics Panel (DREP) under the terms of Anglia Ruskin University's Research Ethics Policy, dated 8th September 2016, Version 1.7, reference FHSE-DREP-17-069. This study was also conducted in accordance with the International Conference for Harmonisation of Good Clinical Practice and the Research Governance Framework for Health and Social Care (Health Research Authority 2016).

4.3.1 Therapists and clinical supervisors

Full, multi-faceted, consideration (Stewart, 2011) was given in relation to the potential impact of this research on its participants. All reasonable attempts were made to minimise the burden on those working in the research setting. As the data collection for this research was routinely gathered as part of normal service delivery, and analysed post hoc, this research did not impact on the therapists or the clinical supervisors who rated the therapists clinical work. The data had already been collected and the transcripts had been rated before this research commenced. Therapists were invited to participate in this research and were asked to provide consent that their anonymised data could be used in order to investigate the therapist variables that might be associated with clinical outcome. Care was taken not to place therapists under any duress to consent for their data to be used. The researcher is aware of the potential power differentiation between herself and the therapists. Therefore, therapists were contacted only once, via email, and were provided with a participant information sheet and consent form. The participant information sheet (see appendix item 7) for this research follows the guidance suggested by Sarantakos (2005) and includes:

- Clear information relating to the researcher
- Clear information about the academic institution and the primary supervisor
- A description about the purpose of the study
- A description relating to the benefits of participating in the study
- Notification relating to the level of participation and the processes involved
- Notification of any risks
- A guarantee relating to confidentiality
- A clear assurance that participation is not obligatory and that participants may withdraw their consent at any time, and this will not be detrimental
- The names and contact details of people who may be contacted if any concerns or questions arise.

No follow-up reminders or chasing emails were sent to the therapists. Care was taken to ensure that the participant information sheet and accompanying email informed therapists that they could withdraw their consent for their data to be used in this research at any time. Withdrawing consent could be done by informing the researcher or a named senior clinician employed by Ieso Digital Health. No therapists withdrew their consent in this research. Therapists were asked if they would like to be kept informed of the findings from this research. Where a therapist had indicated that they wished to be kept informed they have been provided with a summary of the findings and an opportunity to reflect on the potential impact of the findings on their practice with their clinical supervisor. The use of data to inform practice is routine practice in IAPT, and at Ieso Digital Health, therefore this process is unlikely to have been perceived as unusual, threatening or anxiety-provoking. However, in the unlikely event that a therapist was concerned about this research, all therapists were informed that support was available, if required, via any member of the clinical supervisory team, including their own clinical supervisor.

4.3.2 Patients

Ieso Digital Health informs all patients that therapy transcripts and other data will be routinely reviewed by senior clinicians and clinical supervisors for quality control purposes, continuing professional development of the therapist and for research and development. All patients are asked to opt into the use of their data for these purposes. The opt in process is collected and stored electronically using a tick box declaration following guidance and regulations stipulated by NHS England under the auspices of the IAPT programme (NHS Digital, 2019 a,b,c,d,e). Patients are aware that they may withdraw consent at any time and can do so by notifying the Data Protection Officer at Ieso Digital Health or by using the NHS opt out service. No patients withdrew their consent in this research.

IAPT collect anonymised patient data in the form of patient demographics, and outcome data. This data is available in the public domain via the website https:www.nhs.uk. Research and development are considered to be an integral part of IAPT (Clark, 2011) and services are required to request patients opt in for their deidentified data to be used for these purposes (Clark, 2011, NHS Digital, 2019,b,c,d,e). This data is defined as the minimum data set and includes patient gender, age, disability status, geographical location (Clinical Commissioning Group area) ethnicity, outcome measure scores and the primary presenting problem that was treated. The research reported in this thesis used the patient age, gender and PHQ-9 and GAD-7 score at the start of treatment.

4.3.3 Information Security

This research was conducted using therapy transcripts that have derived from online therapy sessions using the Ieso Digital Health platform as a normal and routine part of service delivery. As part of normal practice, the therapy is quality controlled by rating therapists' competence and adherence to NICE and Roth and Pilling (2007,2008) guidelines. Ieso Digital Health follow internationally recognised standards for information security under ISO27001 certification (see https://www.iso.org/isoiec-27001-information-security.html). The company is audited annually in order to retain certification. In addition, Ieso Digital Health are mandated to adhere to the NHS Digital Data Security and Protection Toolkit Standard 2019-20 (see https://www.dsptoolkit.nhs.uk). Ieso Digital Health also has Cyber Essential Plus accreditation (see https://www.cyberessentials.ncsc.gov.uk).

4.3.4 Research design and methods

The researcher is aware of their position of privilege and power, as both an insider researcher and a senior leader in the research setting. Care was taken to consider the impact of herself on the research and on those being researched (the therapists). This involved discussions with therapists and with staff at Ieso Digital Health, in order to ascertain

how this research might be perceived by therapists and how it may impact on them. The ethics of being in a position of trust are not to be understated. Therefore, it was important to ensure that both therapists and staff members were assured that the motives for this research were benevolent and the overarching aims of the research were clear and understandable (Costley, Elliott and Gibbs, 2013). Furthermore, it was important that both participants and staff members felt confident that this research employed sound scientific principles, (Schulz, Altman and Moher, 2010). As the researcher was in a position of power, within the research setting, it was important to ensure that the research design incorporated reasonable steps to minimise the impact of power or influence on those being researched or on those who were otherwise involved in this research. Whilst it is unlikely that all influence was removed, the research design and methods attempted to reduce influence as far as this was possible. These included adopting familiar research methodology, design and methods, provision of inter-rater reliability training, using data that was collected retrospectively (with the exception of the therapist demographic data that was collected after the therapist had consented to participate), providing a clear explanation of the research together with the research question, maintaining a stance of benevolence, resisting attempts to coerce therapists (or other staff) to participate.

Further attempts could have been made to reduce researcher influence in the design in that, for pragmatic reasons, the researcher had rated 33 (16.5%) of the therapists of the therapy transcripts. Albeit that the researcher took care to ensure that the therapists were unknown to her it might have strengthened the design if the researcher has not been involved in rating therapy transcripts, however this would have reduced the sample size to n=167. The researcher decided to proceed with the inclusion of all 200 therapists in that any inferences that were drawn from the statistical analyses could be more reliable. This was particularly important given that the rule of thumb for multivariate analyses require the sample size to 10 times greater than the number of independent variables. It is recognised that the decision to include the researcher ratings in this study may have affected the results.

4.4 METHOD

This research used the data from previously rated therapy sessions and whole treatment episodes (all therapy sessions that were delivered to a patient by the therapist). The practice of reviewing a therapist's clinical work was a normal part of service delivery in the research setting. The overarching aim of this process is to use the data to inform therapists' continuing professional development. The use of live recordings of therapy sessions to support therapists to reflect and learn from their own practice is common practice in CBT (BABCP, 2019), although it is unknown for a service to have access to every therapy session the therapist has delivered.

4.4.1 Therapist demographic data

The n = 200 therapists, who consented to participate in this study, were asked to provide demographic information about themselves. Table 4.1 shows the information that that therapists were asked to provide.

Table 4.1 Questions that were used to collect therapist demographic data

Item	Question	Response
1	Did you attend an IAPT training programme?	Yes/No
2	How would you describe your gender?	Free text
3	To the nearest year how long have you been qualified as a cognitive behavioural therapist?	n
4	What is your core profession?	a) Nurse b) Social Work c) Clinical Psychology d) Counselling Psychology e) None, I used the KSA route to become accredited as a CBT therapist f) Health Psychology g) Other (please specify in the next question)
5	If you answered 'other' to question 3 please specify your core profession	Free text
6	How old are you?	n

The information was initially used to confirm that the therapists were representative of the wider population of High Intensity CBT Therapists. In addition, this data was used as five of the independent variables used to understand whether there is a relationship between any of the demographic variables and clinical outcome.

4.4.2 Rating therapy transcripts

The therapists work included in this research were rated by 6 raters, including the researcher. However, as part of normal service delivery 10 raters had previously rated therapists work. Unfortunately, the intra-class correlation showed that 4 of the 10 raters were consistent outliers, despite inter-rater reliability training (see Chapter 5 for details of intra-class correlation and inter-rater reliability). Inspection of the intra-class correlation showed that one rater had rated 37 of the n=237 therapists who had originally consented. The therapists who had been rated by this rater were therefore excluded from this study.

As part of the routine rating of therapists' sessions, raters were randomly allocated therapists to assess. All of the raters were blind to the clinical outcome and demographic data. Additionally, the raters did not know the therapists they were rating in that they had not previously supervised or taught them. The raters were instructed to randomly select 3 completed cases from the therapist's caseload using the following criteria:

- Each case selected must be a completed case (i.e. the patient must not have dropped out of therapy).
- The case must have been treated at step 3
- The case must have been completed during the time period April 2017 to April 2018.

4.4.3 Rating therapist competence (F score)

Section 2.5.3 of this thesis describes the development of tools that have been used to rate therapist competence, most notably the Cognitive Therapy Scale (CTS) and Cognitive

Therapy Scale-Revised (CTS-R). Other methods of assessing therapist competence have evolved, including the Cognitive Therapy Adherence and Competence Scale (Barber et al., 2003), the Yale Adherence and Competence Scale (Carroll et al., 2000) and, more recently, the Assessment of Core CBT Skills (Muse, McManus, Rakovshik and Thwaites, 2017). None of these tools have been widely adopted in the assessment of CBT therapist competence (Muse et al., 2017) and the CTS-R remains the most commonly used and widely accepted tool of choice to assess therapist competence. Therefore, the CTS-R was selected to rate therapist competence because it is so commonly used. Furthermore, using the CTS-R affords other advantages including the ability compare findings from this research to previous studies and enabling future researchers to undertake similar studies.

As the CTS-R is such a commonly used instrument all the raters in this study had experience of using the CTS-R as a formative and summative assessment tool in IAPT High Intensity CBT training programmes in Higher Education. The raters had also completed inter-rater reliability training together.

Having randomly selected three completed cases from the therapist's caseload (see section 4.4.2, above), the raters were asked to randomly select one therapy session, from each case. The session selected had to be the third, fourth, fifth or sixth treatment session, but not where the sixth session was the final treatment session. Final treatment sessions tend to differ from earlier treatment sessions and often contain less cognitive behavioural interventions. The rater then reviewed the therapy session and scored the session using the CTS-R. This was repeated for each of the three cases, resulting in three scores. The scores were expressed as a percentage. The resulting scores were then recorded as follows:

- Fidelity to the CBT model was recorded as an 'F' score and was entered as the mean of the three CTS-R scores

4.4.4 Rating therapist adherence (A score)

Therapist adherence was assessed using the Roth and Pilling (2007, 2008) competency framework (see Chapter 2, section 2. for further details), National Institute of Health and Social Care Excellence (NICE) guidelines and the evidence base for third wave cognitive behavioural interventions, such as Acceptance and Commitment Therapy (ACT) and Mindfulness Based Cognitive Therapy (MBCT). The raters were asked to record whether, or not, they found sufficient evidence that the therapist was using an evidence-based protocol by recording a mark of 0 where there was insufficient evidence and a mark of 1 where there was evidence that a protocol had been used.

The raters examined 3 completed cases, reviewing every therapy transcript. The raters were asked to decide whether the therapist had used an evidence-based treatment protocol that was appropriate to the patient's primary presenting problem. Adherence to an evidence- based protocol was recorded as an 'A' score where:

A score of 4 indicated that the therapist demonstrated adherence in all 3 cases

A score of 3 indicated that the therapist demonstrated adherence in 2 out of 3 of the cases

A score of 2, indicated that the therapist demonstrated adherence in only one of the cases

A score of 1 indicated that the therapist had not demonstrated adherence in any of the cases.

The resulting score for each therapist were recorded as an A score. Table 4.2 illustrates the complete data set recorded for each therapist.

Table 4.2 Data collected for each therapist

Dependent variable	Independent variables
IAPT clinical outcome definitions: 1. Recovery Rate (expressed as a mean for all patient treated and as a binary outcome) 2. Reliable Improvement (expressed as a mean for all patient treated and as a binary outcome)	1. Demographics: Age, core-profession, training type (IAPT/non IAPT), years of experience and gender 2. F Score 3. A Score 4. The mean score of each of the 12 CTS-R items

4.5 MEASURING OUTCOME

The dependent variables used in the analysis for this study were the IAPT definitions of clinical outcome using the global scores on the PHQ-9 and GAD-7. These measures form part of the IAPT mandatory minimum data set and are routinely used at every appointment within all IAPT services. They were used as a normal part of service delivery and for the post hoc analysis in this research to measure recovery, reliable clinical improvement and minimum clinically significant improvement. A full definition of the terms recovery and reliable improvement can be found in section 2.12 of this thesis.

4.5.1 Outcome measures

The following outcome measures are routinely used at every appointment (see Chapter 2 for full details of the outcome measures). The measures are sent digitally to the patient before each therapy appointment via the Ieso Digital Health platform. The measures are self-administered, and patients complete them via their smart phone, tablet or computer. The outcome measures were used to measure outcome in terms of recovery and reliable improvement.

GAD-7

The GAD-7 (Spitzer et al., 2006) is a seven-item self-report measure for anxiety using a 4-point Likert scale (0-3, where 0 indicates the absence of a symptom and 3 indicates greater severity). A cut off point of ≥ eight indicates greatest sensitivity and a clinical case in a primary care population. A cut of ≥ 15 indicates severe symptoms. The GAD-7 has been demonstrated to be a reliable and valid instrument with a sensitivity of 89%, a specificity of 82% and good internal consistency (Cronbach α =0.92).

PHQ-9

The PHQ-9 (Kroenke, Spitzer and Willams 2001) is a nine-item measure for depression using a 4-point Likert scale (0-3, where 0 indicates the absence of a symptom and 3 indicates greater severity). A cut off point of ≥ ten is used to indicate a diagnosis of depression in a primary care population. A score of ≥ 10 has been demonstrated to be a valid and a reliable tool with a sensitivity of 88% and specificity of 88%, when compared to a mental health professional validation interview (Kroenke, Spitzer and Williams 2001).

Outcome indices

The dependent variable (clinical outcome) was calculated using the IAPT definitions of recovery and reliable improvement (see section 2.12, in this thesis, for a detailed description of the IAPT calculations for both recovery and reliable improvement).

4.6 THERAPIST VARIABLES (INDEPENDENT VARIABLES)

The therapist (independent) variables used in this research fall into three categories; therapist demographic data, therapist competence (F score) and therapist adherence (A score).

4.6.1 Therapists demographics

Therapists provided information about their age, gender, years of experience, type of training they had completed (IAPT course or non IAPT course) and their core profession. In order to become accredited as a CBT therapist, clinicians are required to have a core profession including, registered mental health nurse, doctor, occupational therapist, social worker, clinical psychologist, health psychologist, counselling psychologist, degree in counselling or special educational needs teacher. Clinicians who do not have a core profession are permitted to apply for accreditation via the Key Skills Assessment (KSA) route. The KSA route requires applicants to compile a portfolio of evidence that demonstrates that they have developed the key skills taught on clinical training programmes for the core professions. These skills include topics such as conducting a risk assessment, writing clinical notes, anatomy and physiology, pharmacology and diagnostic criterion. In addition to meeting the KSA criteria at the BABCP require applicants to High Intensity training programmes to have a minimum of two years' experience in a relevant mental health setting (BABCP, 2019).

4.6.2 Therapist competence (F score)

The CTS-R (Cognitive Therapy Scale-Revised) is a twelve item, standardised measure used to assess therapist competence across a range of therapeutic skill areas (Blackburn, Milne and James, Baker, Standart and Garland 2001). Blackburn et al., demonstrated that the CTS-R is a valid and reliable measure of therapist competence and the tool is used to assess how closely a therapist is adhering to the CBT model. The CTS-R is used as both a formative and summative assessment on post-graduate clinical training programmes, for CBT therapists, in the United Kingdom. The instrument has been demonstrated to have high internal consistency (α range = .75–.97 Blackburn et al., Reichelt et al., 2003) and adequate inter-rater reliability (r=0.67). The CTS-R has been widely adopted in United Kingdom and is also commonly used in research studies to demonstrate fidelity to the CBT model (Gordon

2006). The CTS-R assesses a therapist ability across 12 areas (see table 4.3). The first 5 items are general therapeutic skills and the latter 7 items assess CBT specific skills.

Table 4.3 The 12 CTS-R items (Blackburn et al., 2001)

General therapeutic skills	1. Agenda 2. Feedback 3. Collaboration 4. Pacing/use of time 5. Interpersonal effectiveness
CBT specific skills	6. Guided discovery 7. Conceptualisation 8. Identifying key cognitions 9. Application of change methods 10. Application of behavioural techniques 11. Homework 12. Facilitation of emotional expression

Each item is scored on a 7-point Likert scale where a score of 0 indicates an absence of the skill and a score of 6 indicates a high level of expertise has been demonstrated. The total range is 0- 72 with the overall score presented as a percentage. Blackburn et al., provide a training manual, for assessors, detailing example descriptions for each item. A copy of the CTS-R manual can be found at

https://cedar.exeter.ac.uk/iapt/hihandbook/assessments/ctsrmanual/

4.6.3 Therapist Adherence (A) score

Adherence to an evidence-based protocol was assessed using guidelines from three sources Roth and Pilling (2007, 2008) guidelines for the competencies required for the cognitive behavioural treatment of depression and anxiety disorders, NICE guidelines for the treatment of patients with depression or an anxiety disorder and evidence-based treatment protocols documented in CBT academic text books. Raters were asked whether, in their opinion, the therapist had, or had not, demonstrated adherence to an evidence-based

protocol as found in NICE guidelines, Roth and Pilling guidelines or published research cited in CBT academic textbooks.

4.7 INTER-RATER RELIABILITY TRAINING

The raters in this research were required to complete inter-rater reliability training for both assessment of competence, as assessed by the CTS-R, (F score) and adherence (A score). The raters were provided with a CTS-R manual and the guidelines for A score assessments. They were asked to complete F scores and A scores on 4 completed cases. All the raters assessed the F score (competence, as assessed by the CTS-R) on the same transcripts. The raters then attended a whole day training session where they shared their scores with the other raters and provided a rationale for why they had allocated a particular score. The process was repeated with 4 different completed cases and the assessors attended a further day's training.

4.8 PATIENT VARIABLES FOR LOG-LINEAR ANALYSES

Patient demographics and outcome measure data are routinely collected for each patient within IAPT and are commonly used for data analysis and research purposes (Layard and Clark, 2014). Patient variables were used in the final statistical analysis, a log-linear analysis in order to understand whether the predictor variables remained statistically significant. The patient variables which were used in the log-linear analysis were: age, gender and start and end PHQ-9 and GAD-7 scores.

4.9 DATA COLLECTION AND HANDLING

Data was collected and retained in accordance with the Data Protection Act (1998[4]), GDPR, ISO27001 and NHS Data Security Toolkit. Study documents (both paper and electronic)

[4] Correct at the time this research was conducted, although the Data Protection Act (1998) has now been superceded by the Data Protection Act (2018)

were retained in a secure location during and after the study was completed. All documents will be retained for a period of five years.

4.10 STATISTICAL ANALYSES

Tables of descriptive statistics were compiled for all the available variables. For continuous variables various summary statistics were tabulated as appropriate, including the mean, median, standard deviation, interquartile range, and range. For categorical variables counts and percentages were tabulated. Bivariate analysis was conducted in order to investigate the strength of the relationship between the dependent variable (clinical outcome) and each of the independent variables. Pearson's correlation was used for each of the continuous variables and chi square for each of the categorical variables. A linear regression model was used to understand how adherence influences recovery. In order to investigate the nature of the relationship between the dependent variable and the independent variables a multivariate analysis was used, including a multiple regression model and a hierarchical log-linear analysis. A hierarchical log-linear analysis was performed to determine a statistical model for the associations among categorical variables including therapist (F score and A score), patient (age, gender, severity and clinical outcomes (reliable improvement and recovery). Whilst patient variables are not the primary focus of this study the function of adding them to this final analysis was to further examine the therapist A and F score in the context of patient variables. Additionally, the hierarchical loglinear analysis was included in order to explore the differing effects between the variables.

The data was analysed using the software programmes SPSS version 24.0 (IBM, 2016) and JASP (JASP team, 2018).

4.11 CONCLUSION

This chapter presented a rationale for the methodology used for this research, with a review of the theoretical frameworks relevant to the cognitive behavioural model. The chapter went on to outline the research design, methods and statistical analyses used in this research. Chapter 5 will report on the analysis and findings from each statistical procedure.

CHAPTER FIVE: DATA ANALYSIS

Chapter 4 of this thesis outlined the methodology for this research. The unique opportunity to use online therapy transcripts, derived from Internet Enabled CBT (IECBT), to assess therapist competence and therapist adherence was discussed. Chapter 5 will now outline the data analyses and findings from this research. This chapter will begin by defining the descriptive statistics relating to each of the variables. This will be followed by; simple correlations between the predictor variables and the outcome variable and a multiple regression model. Finally, the statistical analysis and findings from a hierarchical loglinear model will be presented.

5.1 INTRODUCTION

The purpose of this study was to understand the relationship between clinical outcome and therapist variables such as age, gender, core profession and type of training (IAPT or non IAPT training programme). It further examined the relationship between clinical outcome and a therapist's ability to deliver cognitive behaviour therapy with fidelity to the model, as measured by the CTS-R, and the therapist's ability to adhere to an evidence-based treatment protocol. This chapter will present the analysis of the collected data commencing with descriptive statistics of each of the variables and moving onto bivariate and multivariate analyses. This remainder of this chapter will report on a series of regression models concluding with a log linear analysis model which will consider the interaction between therapist variables, patient variables and clinical outcome.

5.2 DESCRIPTIVE STATISTICS

The following section will present a summary of the data set used in this research beginning with summary descriptors of the therapist demographics and then proceeding to summarise the data relating to therapist fidelity to the CBT model (F score) and therapist adherence to

an evidence-based protocol (A score) and the outcome variables, recovery rate and reliable improvement rate.

5.2.1 Therapist Demographics

All 200 therapists in this sample provided information regarding their age, gender, years of experience, type of training and core profession. Table 5.1 shows the demographic characteristics of the study sample. Therapist age was captured using a categorical scale, with the majority of therapists aged between 36 and 45. The gender ratio was similar to that observed in general for IAPT services (81.5% female to 18.5% male in this sample vs 82.1% female to 17.9% male for the IAPT therapist workforce). A range of core professions was observed, with the majority of therapists reporting to have no core profession and becoming accredited via the KSA route (45%), followed by Registered Mental Nurses (RMN; 22.5%). In terms of training and experience, the majority of therapists in the sample had been trained on an IAPT training programme (73%), with an overall post-qualification experience ranging from 1 to 20 years (mean = 6.05 years, SD = 3.57 years). Most therapists in this sample had been qualified for between 2 and 10 years (see figure 5.1). This may specifically relate to the development of IAPT in 2009 and the focus on training more CBT therapists since then.

Figure 5.1 Years of post-qualification experience showing that most therapist had been qualified for 11 years or fewer

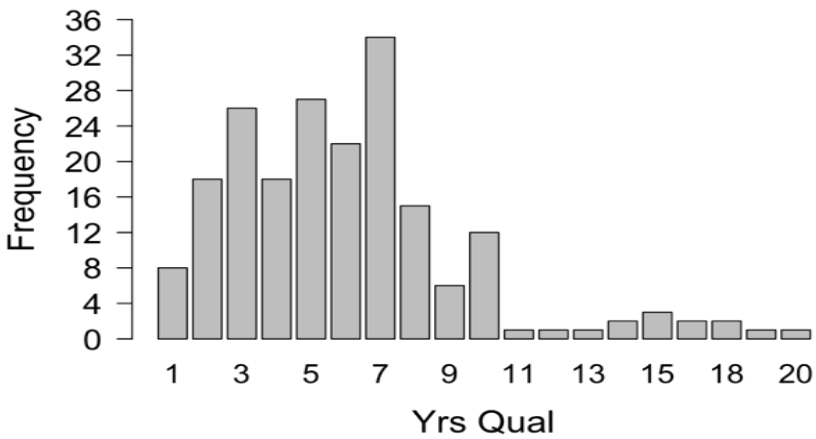


Table 5.1 Demographic characteristics of study sample

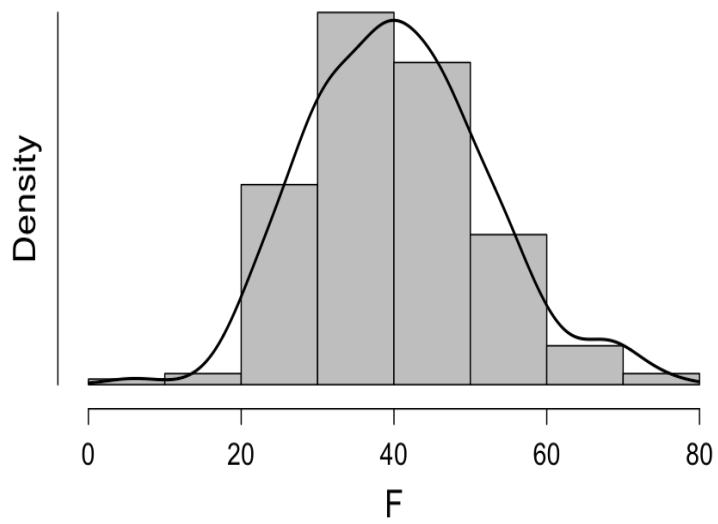
Independent variable	N	Percent
Age		
≤ 25	2	1
26- 35	62	31
36 - 45	76	38
46- 55	38	19
> 55	22	11
Gender		
Male	37	18.5
Female	163	81.5
Core Profession		
Nurse (RMN)	45	22.5
Social worker	8	4
Clinical Psychology	6	3
Counselling psychology	18	9
KSA	90	45
Health Psychology	1	0.5
Other	32	16
Training		
IAPT trained	146	73
Not IAPT trained	54	27

5.2.2 'F scores': fidelity to the model

Therapist competency was measured using the mean of 3 CTS-R's, this is reported as the F score. An F score, (defined as the mean of 3 CTS-Rs), was calculated for all 200 therapists. CTS-Rs were undertaken on randomly selected therapy sessions, excluding the first and last therapy session (see Chapter 3 for further details).

F scores were normally distributed across the sample (mean = 40, SD = 11.65, SE = 0.82, range = 6-75) Figure 5.2 shows the distribution plot for therapist's F scores.

Figure 5.2 Distribution plot for F scores

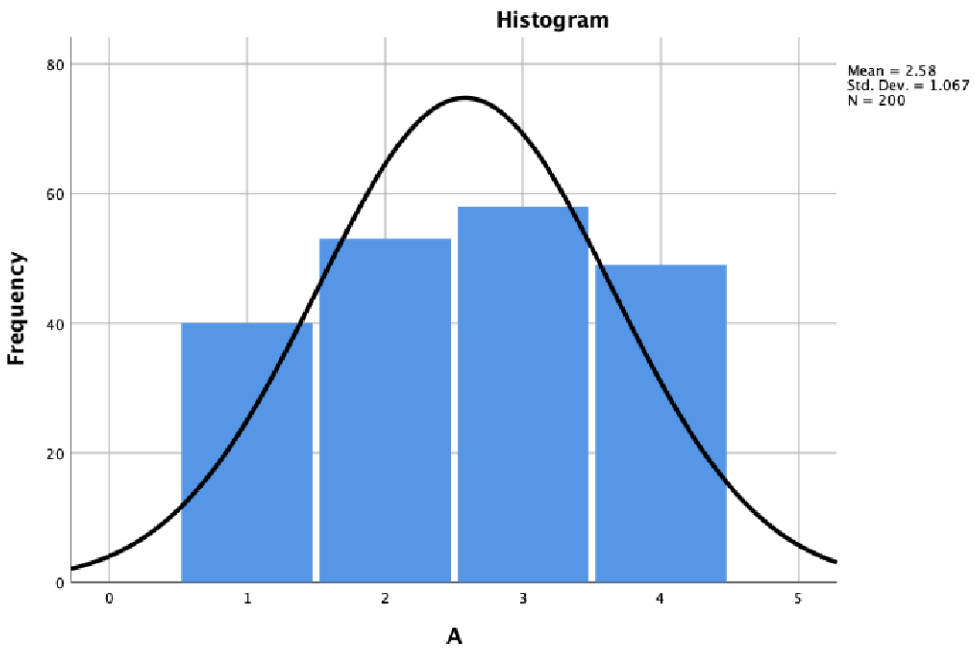


5.2.3 'A scores': adherence to an evidence-based protocol

Therapist adherence to an evidence-based protocol was undertaken by assessing all the therapy transcripts of three completed cases. Therapists' were allocated a total score of between 1 and 4. Where the lowest score of 1 indicates the therapist was not using an evidenced based protocol in any of the cases and the highest score, of 4, indicates that an evidenced based protocol was being used in each case. An A score measuring adherence to an evidence-based protocol was calculated for each of the 200 therapists in the sample.

An A score, ranging from 1 to 4, was assigned to each therapist, where 1 represents a therapist who demonstrated no evidence of a protocol in any of the completed cases assessed (20% of the sample), 2 represents a therapist who demonstrated evidence in one of the three completed cases assessed (26.5% of the sample), 3 represents a therapist who demonstrated evidence in two of the three completed cases assessed (29% of the sample) and 4 represents a therapist who demonstrated adherence to an evidence based protocol in all of the three completed cases assessed (24.5% of the sample). A scores were normally distributed across the sample (mean 2.58, median, 3.00, SD 1.06). Figure 5.3 shows the distribution plot for therapists A scores.

Figure 5.3 Distribution plot for A scores



5.2.4 Patient outcomes - recovery rate and reliable improvement rate

Patient outcomes were recorded using standard IAPT calculations for recovery and reliable improvement (Chapter 4, section 4.5 describes IAPT patient outcome calculations). Each of the therapists in the sample had treated a minimum of 10 patients. The mean recovery rate and reliable improvement rate for 199 of the 200 therapists was calculated. One therapist had missing data. Recovery rate was normally distributed (mean = 50.00%, SD = 19.016%, SE = 1.348%). Reliable improvement rate was normally distributed (mean = 62.98%, SD = 16.08%, SE = 1.14%). Figure 5.4 shows the distribution plot for patient recovery rate and figure 5.5 shows the distribution of patient reliable improvement.

Figure 5.4 Distribution plot for patient recovery rate

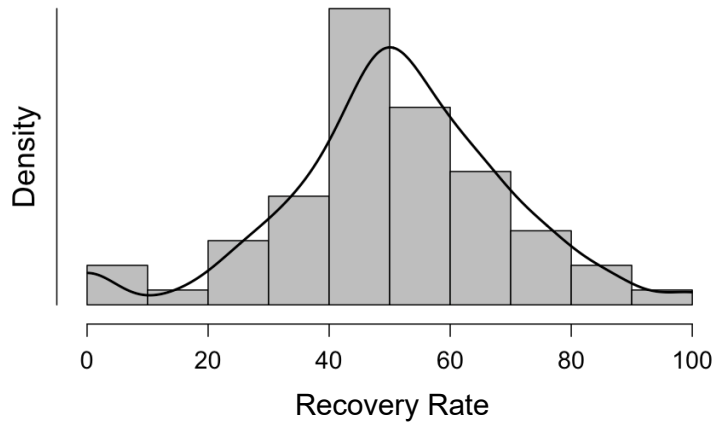
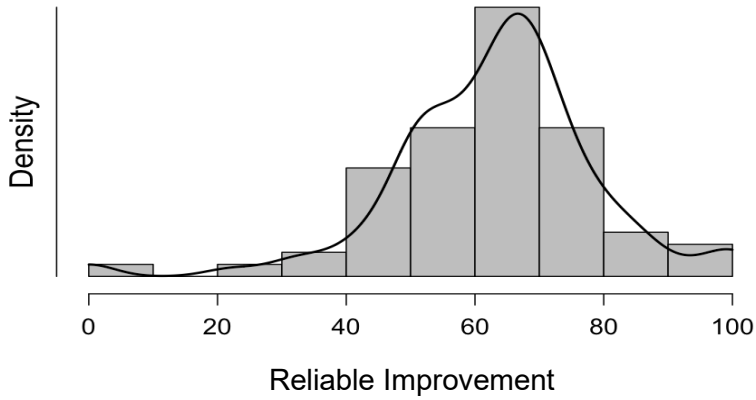


Figure 5.5 Distribution plot for patient reliable improvement rate



5.2.5 Summary of descriptive statistics

This section reported the descriptive statistics for the dependent variable and each of the independent variables. This chapter will now report the inter-rater reliability coefficient for the rating of therapist competence (F score) and therapist adherence to an evidence-based protocol (A score).

5.3 INTER-RATER RELIABILITY

Interrater reliability reflects the variation between 2 or more raters who are measuring the same sample (Koo and Mae, 2016). This study used 6 raters to rate therapist fidelity to the CBT model using the CTS-R and therapist adherence to the evidence base (see chapter four for further details). An Intraclass Correlation Coefficient (ICC) using the McGraw and Wong (1996) convention of two-way mixed effects, absolute agreement with multiple raters was undertaken for both CTS-R rating and rating therapist adherence to the evidence base.

5.3.1 Interrater reliability for CTS-R rating (F score)

The ICC was initially undertaken with 10 raters where a 95% confidence interval of an ICC estimate of 0.716 was 0.359 – 0.944. After further interrater reliability training 4 of the 10 raters were excluded from the study as they were consistent outliers. Using new clinical samples ICC correlation using 6 raters was undertaken where 95% confidence interval of an ICC estimate of 0.983 was 0.946- 0.997.

5.3.2 Inter-rater reliability for rating therapist adherence to the evidence base (A score)

Again, the ICC was initially undertaken with 10 raters where a 95% confidence interval of an ICC estimate of 0.46 was -0.246 – 0.901. After further interrater reliability training the same 4 raters were excluded from the study as they were, again, consistent outliers. Using new

clinical samples ICC correlation using 6 raters was undertaken where a 95% confidence interval of an ICC estimate of 0.953 was 0.857 – 0.992.

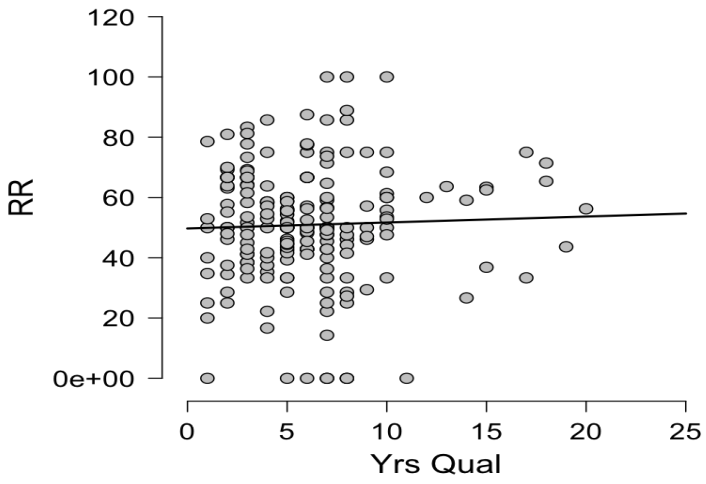
5.4 CORRELATION BETWEEN THERAPIST DEMOGRAPHICS AND OUTCOME

This chapter started with descriptive statistics and inter-rater reliability using the intra-class correlation coefficient. The following section will report on the findings from analysis of the data in order to understand whether any of the predictor variables (therapist demographics, F score and A score) are related to outcome, starting with therapist demographics.

Correlations between therapist variables (age, gender, core profession, experience and training) and recovery rate (dependent variable) were explored using Pearson's correlation and chi-square goodness of fit tests.

For categorical independent variables (age, gender, core profession and training) chi-square tests were performed. The assumptions for a chi square test were met. The data was categorical in nature, contained independence of observations, mutually exclusive and there were at least 5 frequencies in each group. The chi square goodness of fit test revealed no significant associations between any of the variables and recovery rates (age: $X^2 = 1.833$, $p = 0.766$, gender: $X^2 = 0.009$, $p = 0.923$, core profession: $X^2 = 11.521$, $p = 0.074$, training: $X^2 = 96.650$ $p = 0.576$). For the continuous variable (years of experience) a Pearson's correlation was conducted. The assumptions for the test were met. The data was approximately normally distributed. Figure 5.6 shows a scatter plot from a Pearson's correlation between years of experience and recovery rate. The scatter plot illustrates that there was no significant association between years of experience and recovery rate (years of experience $r = 0.037$, $p = 0.602$).

Figure 5.6 Scatter plot of years qualified and recovery rate



5.4.1 Summary

The first question in this thesis queried whether there is a relationship between therapist demographics (age, gender, years of experience and core profession) and recovery rate. This data analysis, using a Pearson's correlation to test the association between therapist years of experience and recovery rate found that there was no statistically significant relationship between the two variables. Similarly, Chi Square tests were conducted to test the association between therapist age, therapist gender, core profession and recovery rate, found that there was no statistically significant relationship between these therapist demographics and recovery rate. It would appear that therapist demographics are not related to recovery rate in this sample.

The second research question in this thesis asked whether there was an association between whether a therapist had completed their CBT training via an IAPT training programme and recovery. A Chi Square test was used to test for an association between the method of training and recovery. No significant relationship was found. It would appear that the method of training (IAPT training programme or other) are not associated with recovery

in this sample. These findings will be discussed more fully in Chapter 6 of this thesis.

5.5 ASSOCIATIONS BETWEEN THERAPIST ADHERENCE (A SCORE), THERAPIST COMPETENCE (F SCORE) AND CLINICAL OUTCOME

Section 5.3 reported on the findings from a Pearson's correlation and chi square tests of association to test the association between therapist demographics (therapist age, gender, core profession, experience, training) and clinical outcome. No statistically significant relationship was found. This chapter will now report on the correlation between therapist adherence (A scores), therapist competence (F scores) and clinical outcome.

5.5.1 Association between therapist adherence (A scores) and clinical outcome

The third and fourth research questions in this thesis query whether there is an association between therapist fidelity to the CBT model and therapist adherence to an evidence-based protocol. In order to answer the fourth question relating to therapist adherence, a linear regression model was used in order to understand how therapist adherence to an evidence-based protocol (A score) influences recovery rate. The assumptions for a linear regression were met. The rule of thumb for sample size (at least 20 cases per independent variable was met. A Durbin-Watson statistic, that assesses independence of residuals, was conducted revealing a Durbin-Watson statistic of 1.840. Visual inspection revealed that the standardised residuals were approximately normally distributed. There was homoscedasticity as assessed by a visual inspection of scatter plot of standardised residuals versus standardised predicted values. The residuals were normally distributed. The descriptive statistics for this model are shown in figures 5.2 (F scores), 5.3 (A scores and 5.4 (recovery) and in sections 5.2.2 (F scores), 5.2.3 (A scores) and 5.2.4 (recovery).

There was a statistically significant relationship between therapist adherence (A scores) and clinical outcome, $r = 0.165$, $p = 0.020$. This would suggest that there is a significant

relationship between therapist adherence and recovery. Table 5.2 shows the model summary and figure 5.7 shows a dot plot and positive regression line for recovery rate. Table 5.2 and Figure 5.7 show that for each increment of the A score recovery increases by 2.9%. A linear regression model was also used to establish whether a similar relationship could be established between the A score and reliable improvement rates. Again, there is a statistically significant relationship between the A score and reliable improvement rate r = 0.182, p = 0.010. Table 5.3 shows the summary model and Figure 5.8 shows a dot plot and positive regression line for reliable improvement rate. Table 5.3 and figure 5.8 show that for each increment of the A score reliable improvement increased by 2.7%.

Table 5.2 Model summary for the linear regression model for A score and recovery rate

Model Summary

Model	R	R^2	Adjusted R^2	RMSE
1	0.165	0.027	0.022	18.803

ANOVA

Model		Sum of Squares	df	Mean Square	F	p
1	Regression	1944.958	1	1944.958	5.501	0.020
	Residual	69652.834	197	353.568		
	Total	71597.792	198			

Coefficients

Model		Unstandardized	Standard Error	Standardized	t	p
1	(Intercept)	43.348	3.492		12.412	< .001
	A	2.931	1.250	0.165	2.345	0.020

Figure 5.7 Scatter plot showing positive regression line for recovery rate

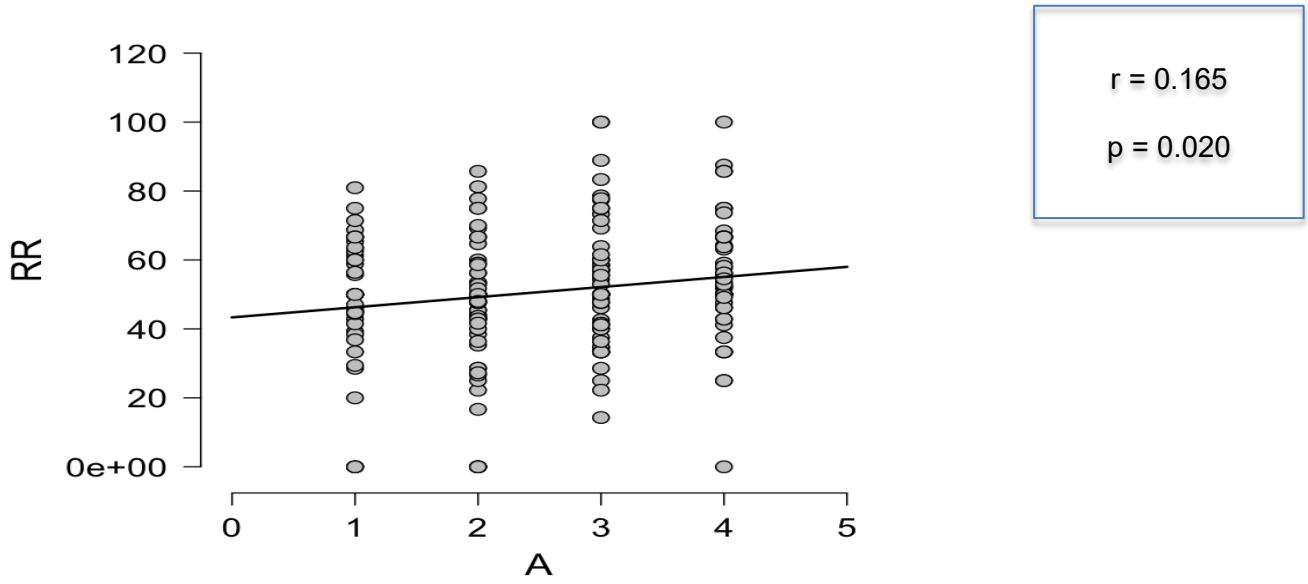


Table 5.3 Model summary for the linear regression model for A scores and reliable improvement rate

Model Summary

Model	R	R^2	Adjusted R^2	RMSE
1	0.182	0.033	0.028	15.855

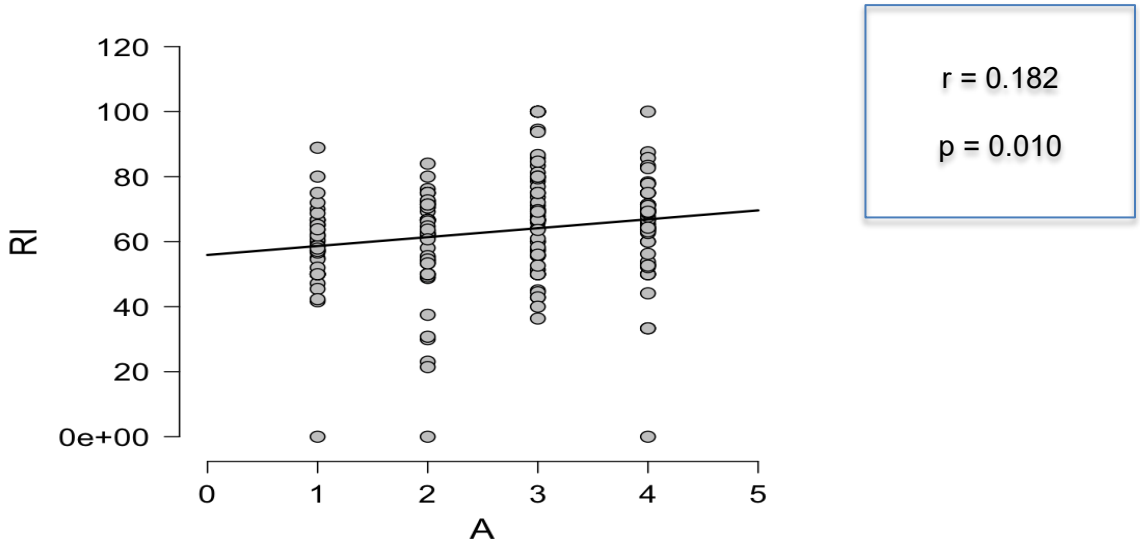
ANOVA

Model		Sum of Squares	df	Mean Square	F	p
1	Regression	1700.914	1	1700.914	6.766	0.010
	Residual	49520.675	197	251.374		
	Total	51221.589	198			

Coefficients

Model		Unstandardized	Standard Error	Standardized	t	p
1	(Intercept)	55.909	2.945		18.986	< .001
	A	2.741	1.054	0.182	2.601	0.010

Figure 5.8 Scatter plot showing positive regression line for reliable improvement rate

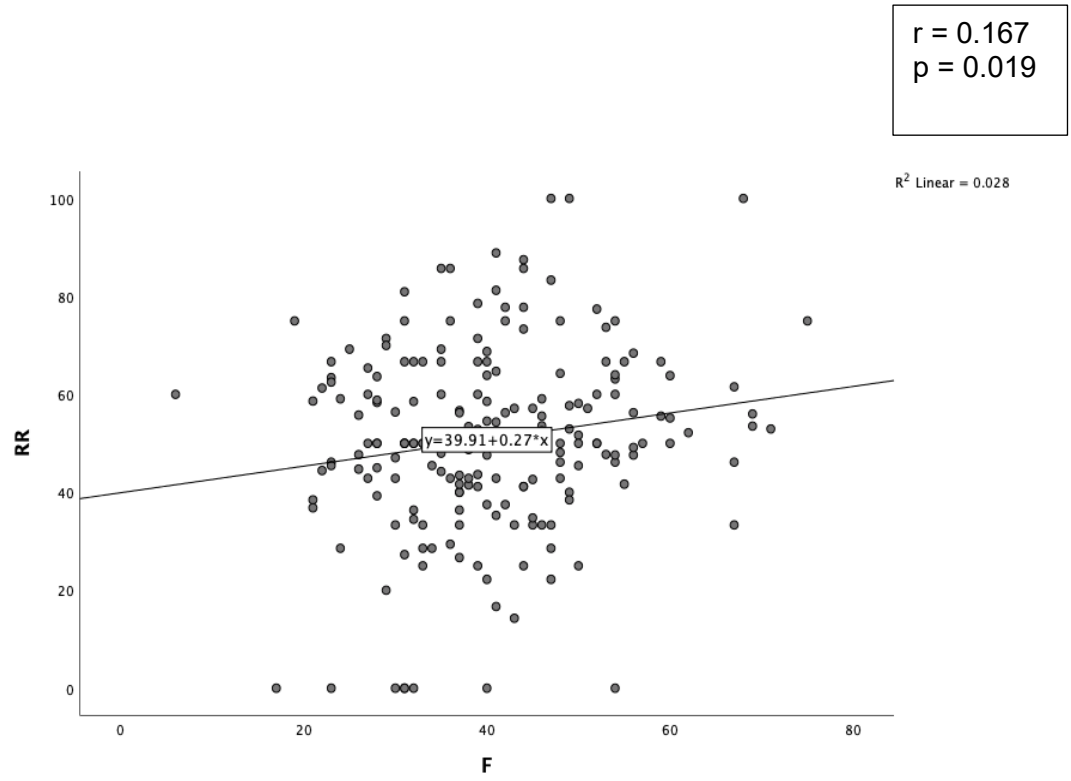


5.5.2 Correlation between therapist fidelity to the model (F scores) and clinical outcome

The third research question in this thesis asks whether there is a relationship between therapist fidelity to the CBT model (F score) and recovery. A Pearson's correlation test was conducted to test the association between the F score and recovery rate. The assumptions for a Pearson's correlation were met. The data were normally distributed.

There was a statistically significant positive correlation between F score and recovery rate, r = 0.167, p= 0.019. This would suggest that there is a relationship between therapist fidelity to the model and recovery. Figure 5.9 shows a scatter plot illustrating the positive correlation.

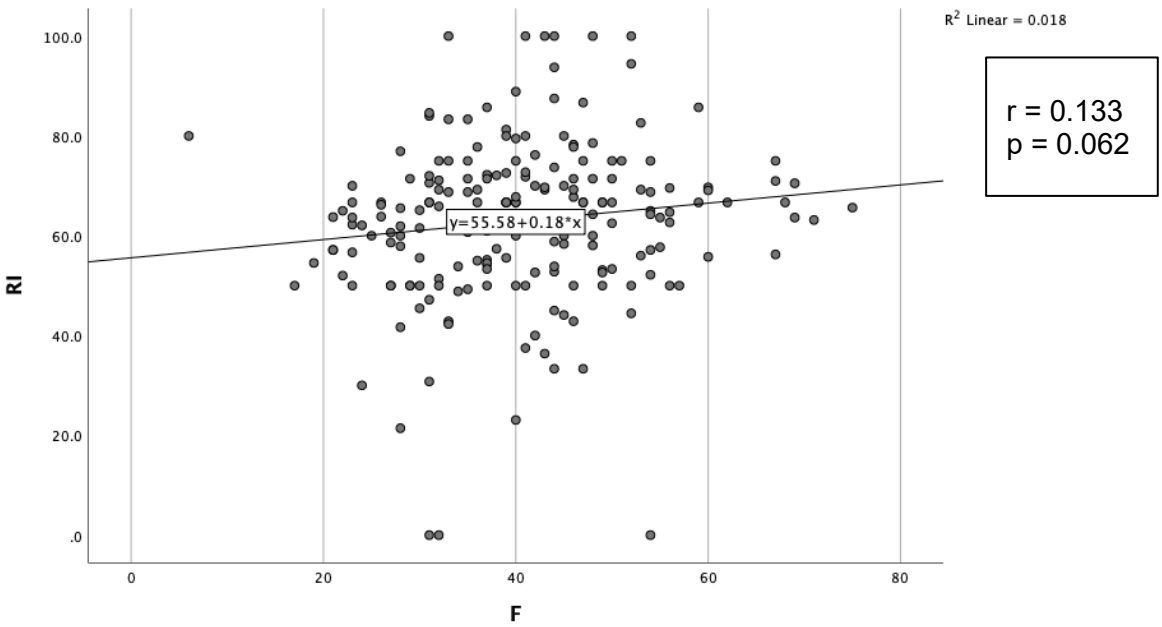
Figure 5.9 Scatter plot of F score and recovery rate



A Pearson's correlation test was also conducted between the F score and reliable improvement rate. The assumptions for a Pearson's correlation were met. Figures 5.2 and 5.4 show that the data were normally distributed.

There was no statistically significant correlation between F score and reliable improvement rate, r = 0.133, p = 0.062. Figure 5.10 shows a scatter plot illustrating the correlation between F scores and reliable improvement rate.

Figure 5.10 Scatter plot of F score and reliable improvement rate



5.5.3 Testing for a relationship between individual CTS-R scores and recovery rate

Section 5.4.2 discussed the findings from a Pearson's correlation test to test the association between therapist fidelity to the CBT model and recovery. The findings would suggest that therapist F scores have a statistically significant relationship with recovery rate ($r = 0.167$, $p = 0.019$). As this was a significant finding a linear regression model was used to further explore the relationship between the mean score of each of the individual items of the CTS-R and recovery rate, in order to understand whether higher scores on individual items of the CTS-R might predict recovery.

To test for the assumptions of a linear regression analysis a Durbin-Watson statistic, that assesses independence of residuals, was conducted revealing a Durbin-Watson statistic of 1.873. Visual inspection revealed that the standardised residuals were approximately normally distributed. There was homoscedasticity as assessed by a visual inspection of a

scatterplot of standardised residuals versus standardised predicted values. The residuals were normally distributed as shown in the standardised residuals plot in figure 5.11.

Figure 5.11 Histogram showing normal distributed standardised residuals

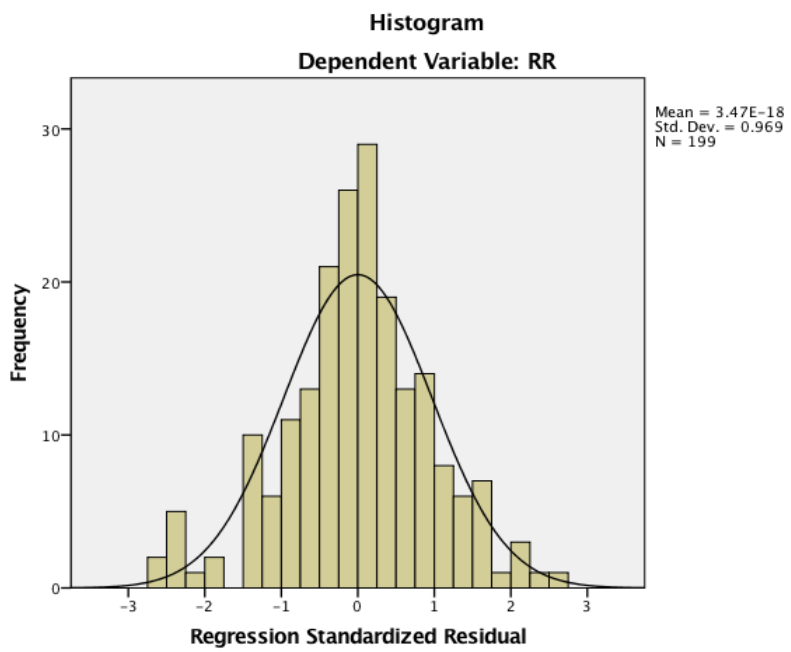


Table 5.4 The Linear Regression Model

ANOVAa

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	4910.375	12	409.198	1.141	.329[b]
	Residual	66687.417	186	358.534		
	Total	71597.792	198			

a. Dependent Variable: RR

b. Predictors: (Constant), Mean of each CTS-R item

The linear regression model is shown in table 5.4. The model failed to detect a statistically significant effect of any single CTS-R item on clinical outcome, using recovery as the dependent variable. The coefficients from the model are reported in table 5.5 showing that none of the individual items were significant.

Table 5.5 Coefficients from linear regression model showing no significant findings

CTS-R item	B	SE	t	Sig.
1	-.993	1.597	-.622	.535
2	5.758	3.465	1.662	.098
3	-4.052	3.756	-1.079	.282
4	-2.216	3.086	-.718	.474
5	-1.866	3.462	-.539	.590
6	2.464	2.961	.832	.406
7	.376	3.045	.124	.902
8	.594	2.761	.215	.830
9	5.248	3.198	1.641	.103
10	.148	2.095	.071	.944
11	-.512	2.720	-.188	.851
12	-.949	2.706	-.351	.726

5.5.4 Summary

This section has reported on the correlation between therapist fidelity to the CBT model (F score) and patient outcome using both recovery rate and reliable improvement rate. The relationship between F score and recovery rate was significant at the < 0.05 level. However, the relationship between F scores and reliable improvement was not statistically significant. As the relationship between F scores and recovery rate was significant this section reported on a secondary analysis that used a linear regression model to predict whether individual CTS-R scores might predict outcome. The findings from this model indicated that no individual CTS-R item was significantly associated with outcome, when adjusted for the other 11 items.

Sections 5.4.1 and 5.4.2 in this chapter reported that A and F scores were correlated with recovery. Given that these were both significant findings, this chapter will now discuss the analyses and findings from a Multiple Regression Model used to explore the effect of A and F scores in the context of the other independent variables.

5.6 MULTIPLE REGRESSION MODEL: F SCORES AND A SCORES

This chapter has so far reported that individual correlation analyses revealed significant positive associations between therapist fidelity to the CBT model (F score), therapist adherence to an evidence-based protocol (A score) and clinical outcomes as measured by recovery rate. However, these analyses fail to explore the effect of these variables beyond the variance that can be explained by other variables such as, for example, therapist demographic variables. For this reason, a multiple regression model was conducted to explore the effect of F and A scores on clinical outcomes in the context of other measured variables. This section will report on how this model was used to further explore the relationship between the F score, A score and recovery. The model summary, as seen in table 5.6, sequentially adds independent variables starting with A scores and F scores,

allowing for the control of each independent variable, in order to answer the supplementary question; how much variation can be accounted for by sequentially adding additional independent variables? The descriptive statistics for this model are shown in figures 5.2 (F scores), 5.3 (A scores and 5.4 (recovery) and in sections 5.2.2 (F scores), 5.2.3 (A scores) and 5.2.4 (recovery).

A multiple linear regression was conducted. There was linearity, as assessed by partial regression plots and a plot of studentised residuals against the predicted values. There was independence of residuals, as assessed by a Durbin-Watson statistic of 1.872. There was homoscedasticity, as assessed by a visual inspection of a plot of studentised residuals versus unstandardized predicted values. There was no evidence of multicollinearity, as assessed by tolerance values greater than 0.1. There were no studentised deleted residuals greater than ± 3 standard deviations, no leverage values greater than 0.2 and values for Cook's distance above 1. The assumption of normality was met following inspection of the Q-Q plot.

Table 5.6 Model summary of the Multiple Regression Model used

					Change Statistics					
Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	R Square Change	F Change	df1	df2	Sig. F Change	Durbin-Watson
1	.188[a]	.035	.025	18.79810429328	.035	3.572	2	195	.030	
2	.205[b]	.042	.027	18.77939737325	.007	1.389	1	194	.240	
3	.206[c]	.043	.023	18.82388998766	.000	.084	1	193	.772	
4	.207[d]	.043	.008	18.96910563369	.000	.019	3	190	.996	1.871

a. Predictors: (Constant), A, F

b. Predictors: (Constant), A, F, YrsQual

c. Predictors: (Constant), A, F, YrsQual, Profession

d. Predictors: (Constant), A, F, YrsQual, Profession, Gender, IAPT, Age

e. Dependent Variable: RR

Tables 5.7 and 5.8 show the sequential results from the hierarchical regression model where model 1 contains only the A score and F score variables, in model 2 the variable therapist years qualified was added, model 3 therapist profession was added and model 4 all the remaining variables were added. Table 5.7 shows the F test statistic resulting from the model. Whilst model 1 (A scores and F scores) is significant (F= 3.572, p= .030), the full model, which includes all the variables is insignificant (F= 1.216, p= .296). Table 5.8 shows the coefficients, confidence intervals and t test results from the model. The table shows that no other variable was found to be significant when fitted after others. This finding is unsurprising given that none of the demographic variables (years qualified, profession, age, gender and training) were found to be significant in themselves in the earlier analyses reported in this chapter.

Table 5.7 Sequential results from the F test in the Multiple Regression Model showing that whilst model 1 is significant (p= .030) the total model with all the variables added is insignificant (p=. 296).

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	2524.209	2	1262.104	3.572	.030[b]
	Residual	68906.901	195	353.369		
	Total	71431.110	197			
2	Regression	3013.952	3	1004.651	2.849	.039[c]
	Residual	68417.159	194	352.666		
	Total	71431.110	197			
3	Regression	3043.715	4	760.929	2.147	.076[d]
	Residual	68387.395	193	354.339		
	Total	71431.110	197			
4	Regression	3063.986	7	437.712	1.216	.296[e]
	Residual	68367.124	190	359.827		
	Total	71431.110	197			

a. Dependent Variable: RR

b. Predictors: (Constant), A, F

c. Predictors: (Constant), A, F, YrsQual

d. Predictors: (Constant), A, F, YrsQual, Profession

e. Predictors: (Constant), A, F, YrsQual, Profession, Gender, IAPT, Age

Table 5.8 Coefficients, confidence intervals and t test scores from the Multiple Regression Model showing insignificant findings

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	95.0% Confidence Interval for B		Correlations			Collinearity Statistics	
		B	Std. Error	Beta			Lower Bound	Upper Bound	Zero-order	Partial	Part	Tolerance	VIF
1	(Constant)	39.163	4.843		8.086	.000	29.611	48.715					
	F	.177	.156	.109	1.138	.257	-.130	.484	.174	.081	.080	.542	1.845
	A	1.715	1.698	.096	1.010	.314	-1.634	5.065	.170	.072	.071	.542	1.845
2	(Constant)	35.474	5.763		6.155	.000	24.107	46.840					
	F	.179	.156	.110	1.151	.251	-.128	.486	.174	.082	.081	.542	1.846
	A	2.046	1.720	.115	1.190	.236	-1.346	5.439	.170	.085	.084	.527	1.896
	YrsQual	.453	.385	.085	1.178	.240	-.305	1.212	.041	.084	.083	.948	1.055
3	(Constant)	36.193	6.287		5.757	.000	23.792	48.593					
	F	.183	.156	.112	1.169	.244	-.126	.491	.174	.084	.082	.538	1.859
	A	2.033	1.725	.114	1.179	.240	-1.369	5.434	.170	.085	.083	.527	1.897
	YrsQual	.448	.386	.084	1.160	.248	-.314	1.209	.041	.083	.082	.946	1.058
	Profession	-.193	.666	-.021	-.290	.772	-1.506	1.120	-.009	-.021	-.020	.988	1.012
4	(Constant)	36.284	10.211		3.554	.000	16.144	56.425					
	F	.180	.159	.111	1.134	.258	-.133	.494	.174	.082	.080	.529	1.891
	A	2.059	1.742	.116	1.181	.239	-1.378	5.496	.170	.085	.084	.524	1.907
	YrsQual	.436	.419	.082	1.039	.300	-.391	1.263	.041	.075	.074	.814	1.229
	Profession	-.176	.707	-.019	-.249	.804	-1.571	1.219	-.009	-.018	-.018	.889	1.125
	Age	.210	1.539	.011	.136	.892	-2.826	3.245	.005	.010	.010	.790	1.266
	Gender	-.668	3.537	-.014	-.189	.850	-7.645	6.309	-.025	-.014	-.013	.956	1.046
	IAPT	.071	3.158	.002	.023	.982	-6.159	6.302	.013	.002	.002	.918	1.089

a. Dependent Variable: RR

5.6.1 Summary of Multiple Regression Model

The full model of therapist age, gender, core profession, training, experience, F score and A score was not statistically significant $R^2 = 0.043$, $F = 1.216$, $p = 0.296$. However, table 5.7 shows that model 1 (independent variables = F scores and A scores) is statistically significant at the 0.05 level ($p= 0.030$). However, in models 2, 3 and 4 the sequential addition of the other independent variables result in $<$ 1% change between the models and models 2, 3 and 4 are not statistically significant with; model 2 R^2 change = .007, $p= 0.240$, model 3 R^2 change = .000, $p= 0.772$ and model 4, R^2 change = 0.000, $p= 0.996$. Beta coefficients (see table 5.8) for the seven independent variables were : A score, $\beta = .116$, $t = 1.181$, $p = .239$, F score, $\beta =.111$, $t = 1.134$, $p= .258$, years qualified $\beta = .082$ $t = 1.039$ $p = .300$, profession $\beta = -0.19$, $t = -.249$, $p = .804$, age $\beta = .011$, $t = .136$, $p = .892$, gender $\beta = -0.14$, $t = -.189$, $p = .850$, IAPT training, $\beta = .002$, $t = 0.23$, $p = .982$. The best fitting model for predicting recovery rate is a linear combination of A scores and F scores. Additions of the other independent variables did not significantly improve prediction.

Given the findings from this model, a final model was used to determine the differing effects between patient outcomes, therapist demographics, therapist F score and A score and three new variables; patient age, patient severity on PHQ-9 and GAD-7 and patient gender. The final section in this chapter will present the analyses and findings from a hierarchical loglinear analysis. This final analysis was conducted in order to determine a statistical model for the associations amongst all the therapist variables and clinical outcomes (improvement, recovery). Unlike the other analyses this model adds patient variables (age, severity and gender) to the model in order to understand the differing effects between all the variables. Log-linear models offer a valuable systematic approach when analysing complex multi-dimensional contingency tables, allowing comparative analyses of differing effects between variables to be undertaken (Everitt, 1977).

5.7 HIERARCHICAL LOGLINEAR ANALYSIS

The final section in this chapter will present the analysis and findings from a hierarchical loglinear analysis. The third and fourth research questions in this thesis ask whether there is a relationship between therapist fidelity to the CBT model (F score) and clinical outcome and whether there is a relationship between therapist adherence and clinical outcome. Unlike the other analyses described in this chapter, this analysis will explore the associations between therapist F score/A score and outcome in the context of therapist demographics and three new variables, patient age, patient gender and patient severity, at the start of treatment, on the PHQ-9 and GAD-7. Whilst patient variables are not the focus of this research, adding them to this final analysis further explores whether an association between therapist F and A scores and outcome remains when new variables are added to the model. Loglinear models offer a valuable systematic approach when analysing complex data, allowing comparative analyses of differing effects between variables to be undertaken (Everitt, 1992). Whilst loglinear models are a novel application in psychological therapy research they are commonly used medical research (see Maimaris et al.,1994, Helmy et al., 2010). Hierarchical loglinear models do not discriminate between dependent and independent variables. The model treats all variables as equal, thus making it possible to see which variables are associated with each other and which are not (Everitt, 1992). Unlike other statistical analyses (such as logistic regression) that distinguish between dependent and independent variables the approach adopted has the advantages of considering all variables on an equal footing for data exploration purposes. Hierarchical loglinear models do not deal with nested variables nor do they produce directions between factors. Hierarchical loglinear models do, however, serve a different purpose primarily for complex data exploration and understanding by reducing high dimensional, cross-classified categorical data into a format that captures the likely generalisable features of the particular multivariate data set sampled in a way that readily allows the creation of a visual map of

significant interdependencies among the variables. For this reason, hierarchical loglinear models are sometimes known as graphical loglinear models.

The model requires that all variables are converted to categorical variables. Table 5.9 shows how each variable was converted using a median split. The median for patient age, and severity on PHQ-9 and GAD-7 is shown in table 5.10.

Table 5.9 Conversion to categorical variables

Patient variables	**Conversion to categorical data**
Patient age	0/1 (Older or younger than the median)
Patient gender	0/1 (male/female)
Severity PHQ-9	0/1 (Above or below the median)
Severity GAD-7	0/1 (Above or below the median)
Recovery	0/1 (yes/no)
Reliable improvement	0/1 (yes/no)
Therapist variables	
Adherence to an evidence-based protocol	0/1 (yes/no for each patient treated)
F score	1-4 based on which quartile the F score fell within

5.7.1 Descriptive statistics for all categorical variables in the model

594 patients were included in the model, 6 patients had missing data. 53.3% were younger than the median age (33 years) and 46.7% were equal to or older than the median age (median =33, mean = 35.08, standard deviation 11.92). 73.8% were female and 26.2% were male. 54.% scored above the median for severity, at the start of treatment, on the PHQ-9 (median = 13.0, mean = 13.41, standard deviation = 6.18) and 55.9% scored above the

median for severity, at the start of treatment on the GAD-7 (median = 13.0, mean = 12.82, standard deviation = 5.15). 75.9% reliably improved and 54.9% recovered. Table 5.10 shows the frequencies for each variable in the model.

Table 5.10 Frequency table for patient and therapist variables showing median, mean and standard deviation for patient age and severity (PHQ-9 and GAD-7)

Patient variables	Percentage %	Median	Mean	SD
Age: Younger	53.3	33	35.08	11.92
Age: Older	46.7			
Gender: Male	26.2			
Gender: Female	73.8			
Severity: PHQ-9 < median	45.5	13.0	13.41	6.18
Severity PHQ-9 ≥ median	54.5			
Severity: GAD-7 < median	44.1	13.0	12.82	5.14
Severity GAD-7 ≥ median	55.9			
Rel. Improv. Yes	75.9			
Rel. Improv. No	24.1			
Recovery Yes	54.9			
Recovery No	45.1			

5.7.2 The hierarchical loglinear analysis model

A hierarchical loglinear analysis was performed to determine a statistical model for the associations among categorical variables including therapist (fidelity score and adherence score), patient (age, gender and initial anxiety and depression severity scores) and clinical outcomes (improvement, recovery). Loglinear models offer a valuable systematic approach when analysing complex multi-dimensional contingency tables, allowing comparative analyses of differing effects between the variables to be undertaken (Everitt, 1977). In this analysis, 594 cases were included in the model, derived by a backwards elimination procedure beginning with an unsaturated model for all eight main effects and their 28 possible two-way interactions. The final model was derived using SPSS (hierarchical loglinear model selection, IBM SPSS, version 24). The assumptions for the model were met in that there were no outliers and approximately normally distributed residuals. The SPSS

output from the model is shown in appendix item 8. This computer-generated model included gender as main effect only and the two-way interactions between A score*F score, age*GAD-7 at start of treatment, age* recovery, F score*recovery, GAD-7*improvement, GAD-7*PHQ-9, GAD-7*recovery. PHQ-9*improvement and improvement*recovery, having excluded all other interactions as non-statistically significant. The model had a likelihood ratio of X^2 (488) = 405.3, p= .997. A graphical representation of this model, indicating the relative strengths of the various interactions, is shown in figure 5.12 with the generating class shown in table 5.11.

Table 5.11 Generating class for best hierarchical log linear model

Interaction	X^2	df	p value
A*F Quartile	217.9	3	<0.0005
Age*GAD	6.7	1	.010
Age*Recovery	4.9	1	.025
F Quartile *Recovery	10.1	3	.018
GAD*Improvement	27.3	1	<0.0005
GAD*PHQ	100.8	1	<0.0005
GAD* Recovery	13.1	1	<0.0005
PHQ*improvement	10.4	1	.001
Improved *Recovery	183.7	1	<0.0005
Gender	139.4	1	<0.0005

The model includes 28 possible interactions. Figure 5.12 shows the ‘saturated model’ with all 28 interactions, meaning such a model would fit the data perfectly. In standard hierarchical loglinear model notation this corresponds to the model [DFRGADPI] with factors abbreviated to their initial letters, except “D“ for GAD and “H” for Adherence. For further details on methodology, including notation, see Gauraha (2017). It should be noted that the saturated model is used for baseline comparison purposes when assessing more models containing fewer terms. The usefulness of hierarchical loglinear models is to see if a more useful and plausible model can emerge that only retains the significant interdependencies among the variables. The final “fitted model”, depicted in figure 5.13, is generated by use of

SPSS routine iterating backwards elimination of the non-significant interaction terms in the corresponding models.

Figure 5.12 Loglinear analysis ‘saturated model’ showing all 28 possible interactions

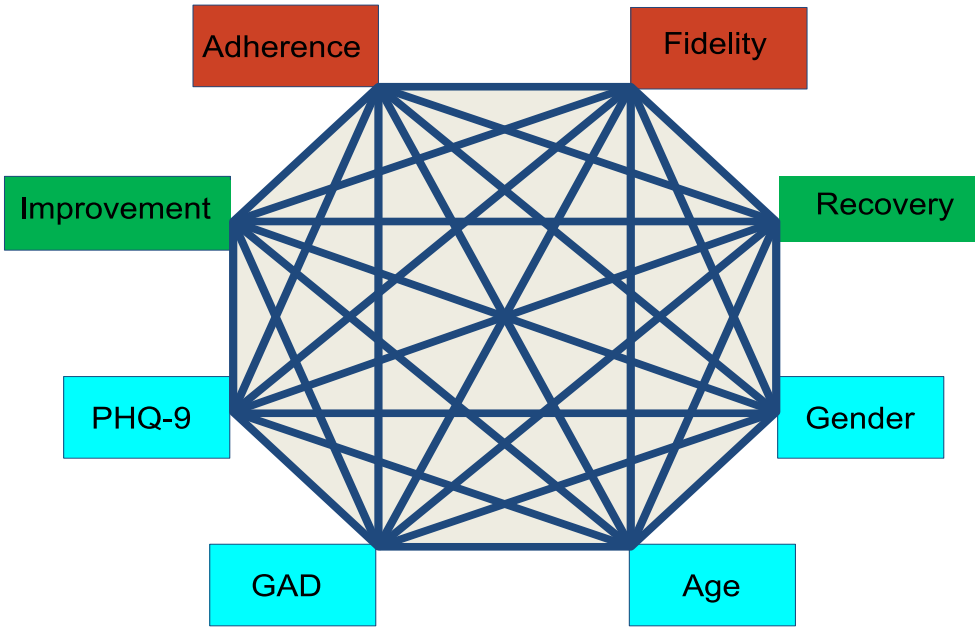


Figure 5.13 The Hierarchical Loglinear 'fitted model' showing how the model eliminates non-significant links, one at a time until only significant links remain**.**

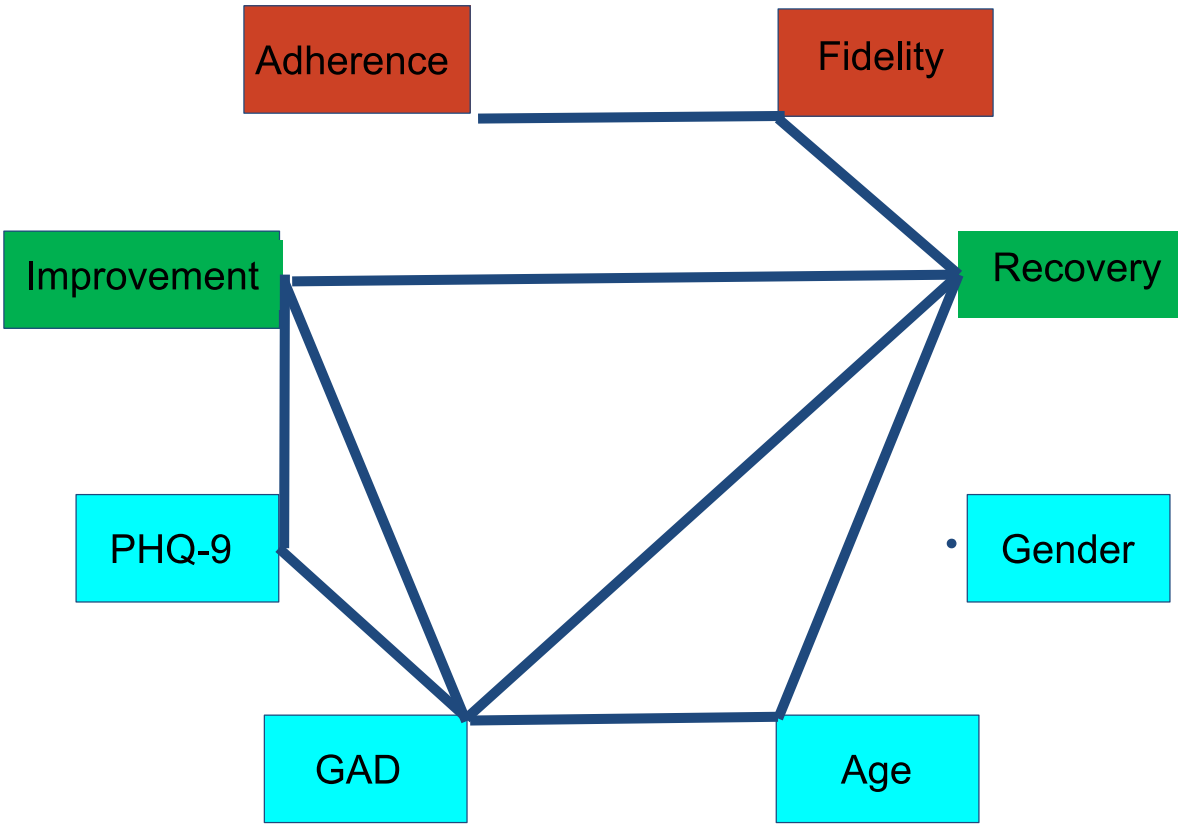


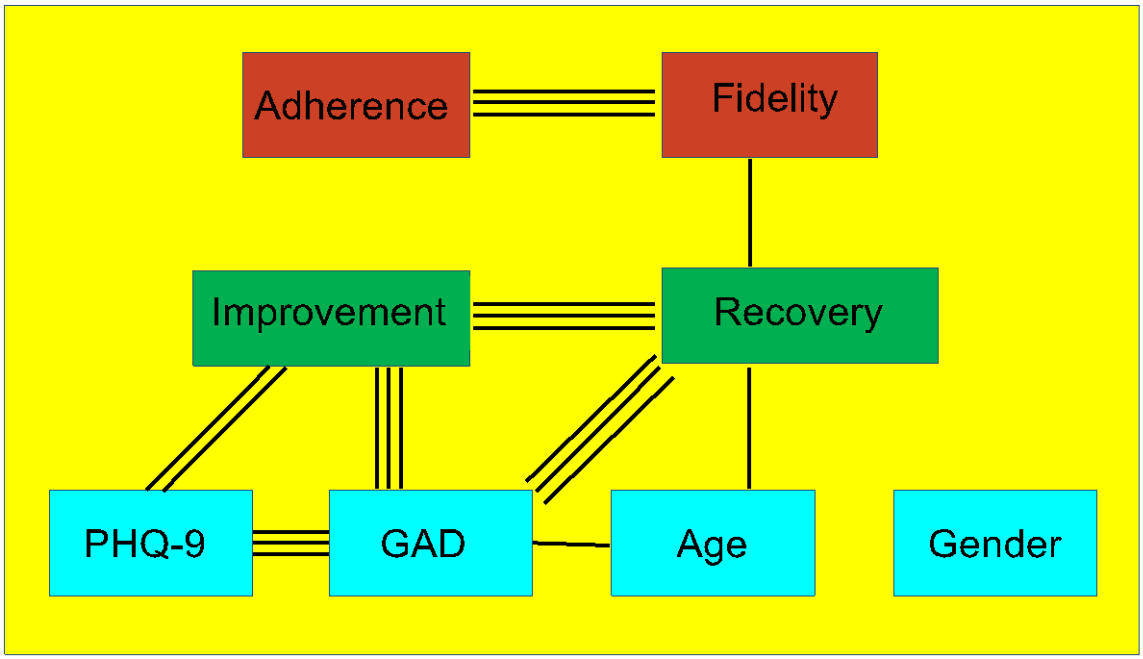
Figure 5.13 shows how the model eliminates all non-significant links until only significant links remain. Once no more pairwise links can be deleted without statistically significant loss to the resulting model the automated procedure generates a final 'fitted model' depicted in figure 5.13 and represented notationally as: [G] [HF] [FR] [RAD] [RID] [IPD]. Therefore, the final model corresponds to having one main effect for gender, two two-way interactions and three three-way interaction terms. The graph, as shown in figure 5.13 maps the data for relative ease of visual interpretation. Figure 5.14 shows the same model with some repositioning and colour-coding for added clarity. Figure 5.14 also indicates the relative strengths of the links, with triple lines representing the strongest. Whether links are weak, moderate or strong is determined by the order of magnitude of statistical significance from the SPSS output for the final model. It should be noted that where links are present the direction of influence is not inherently produced in the modelling, merely the presence or absence of significant interactions in the overall model.

5.7.3 Interpreting the loglinear analysis model

The final model, as shown figure 5.14, shows the associations between the variables where a single line between 2 variables indicates $p = <0.05$, two lines indicates $p = <0.01$ and 3 lines indicates $p = <0.001$. Figure 5.14 shows that, in therapist variables, A scores and F scores significantly interacted ($p = <0.001$) and F scores interacted with recovery ($p = <0.05$). Recovery, in turn, was significantly related to the GAD-7 score ($p = <0.001$). The GAD-7 score was the pre-intervention score and was used as an indicator of severity at the start of treatment. The GAD-7 and PHQ-9 scores were also significantly related ($p = <0.001$), suggesting that patients were likely to be severe/not severe on both outcome measures at the start of treatment. This is a reassuring finding in that it corroborates the model given that there is existing evidence that severity on the GAD-7 and PHQ-9 are related (Catarino et al., 2016). The other patient variable related to the outcome variable of recovery was patient age ($p = <0.05$). Additionally, age was related to GAD-7 scores ($p = <0.01$) suggesting that older

patients, in this sample, were more likely to present with anxiety disorders and were also more likely to recover. Patient gender was important in this model only in as much as it was a main effect in the model, but it did not significantly interact with any of the other variables. Table 5.12 shows the associated estimated odds ratio. These are used to quantify size of effects, whether above or below one indicates the direction of association. It should be noted that of the 12 odds ratios tabulated all but two have confidence intervals spanning the neutral value 1, with GAD-7 and PHQ-9 scores, respectively, being statistically significantly related to improvement. The HLM analysis above, however, provides a more satisfactory, overall picture of the inter-relationships among all the factors simultaneously.

Figure 5.14 The final model showing associations between the variables



Therapist variables

Outcome variables

Patient Variables

p = <0.05

p = <0.01

p= <0.001

Table 5.12 Odds ratios (ORs) and 95% confidence intervals (CIs) for reliable improvement and clinical recovery by patient factors (gender; age below/above median; low/high GAD and PHQ scores) and therapist factors (Adherence score low/high; Fidelity score low/high).

		Improvement		Recovery	
		n (%)	OR (95% CI)	n (%)	OR (95% CI)
Gender	Male	116 (74.8%)	0.93 (0.61, 1.42)	87 (56.1%)	1.07 (0.74, 1.54)
	Female	334 (76.3%)		239 (54.6%)	
Age	< 33	246 (77.1%)	1.25 (0.86, 1.82)	187 (59.2%)	1.45 (1.05, 2.01)
	33+	205 (73.7%)		139 (50.0%)	
GAD-7	Low	169 (64.5%)	0.32 (0.22, 0.48)	145 (53.3%)	1.03 (0.75, 1.43)
	High	282 (84.9%)		181 (54.5%)	
PHQ-9	Low	177 (65.6%)	0.35 (0.24, 0.51)	146 (54.1%)	0.94 (0.68, 1.30)
	High	274 (84.6%)		180 (55.6%)	
Adherence	Low	215 (73.6%)	0.78 (0.54, 1.14)	150 (51.4%)	0.76 (0.55, 1.05)
	High	236 (78.2%)		176 (58.3%)	
Fidelity	Low	208 (72.5%)	0.69 (0.48, 1.01)	146 (50.9%)	0.73 (0.53, 1.01)
	High	243 (79.2%)		180 (58.6%)	

5.7.4 Summary

This final analyses in this chapter have shown that there is a significant relationship between therapist competence, as measured by the CTS-R, and recovery. Therapist adherence is strongly associated with therapist competence and, therefore, adherence is associated with recovery but only through competence. These findings support those that have been reported previously in this chapter in that therapist competence is associated with recovery. The addition of patient variables to the model has highlighted the interaction between patient and therapist variables and the findings have demonstrated that competence and adherence remain significant when patient variables are added to the model. It should be emphasised that just as a correlation analysis, between a pair of variables (e.g. with correlation coefficient numerically close to 1) does not imply causation, so too is this model-building and model-displaying tool best seen as an exploratory analysis, not one for establishing causal links.

5.8 CONCLUSION

This chapter has presented the findings from the data analysis used to understand the relationship between the independent variables of therapist age, therapist gender, therapist length of experience, therapist core profession, therapist training, therapist competence, therapist adherence and, the dependent variable, clinical outcome. The key findings from this chapter are that therapist age, gender, core profession, training and length of experience are not related to clinical outcome. However, therapist adherence (reported as an A score) has a statistically significant relationship at the < 0.05 level. Similarly, therapist competence (reported as an F score), as measured by the CTS-R, has a statistically significant relationship at the < 0.05 level with item 2 of the CTS-R being significant at the < 0.01 level and items 6, 7, 8 and 9 being significant at the < 0.05 level. In addition, the relationship between F scores (therapist competence as measured by the mean of 3 CTS-R scores) and A scores (therapist adherence) and clinical outcome remain significant when

placed into a hierarchical linear regression model. Furthermore, these findings remain significant using a loglinear analysis model when patient variables are added to the model. These results will be discussed and expanded upon in the following chapter.

CHAPTER SIX: DISCUSSION

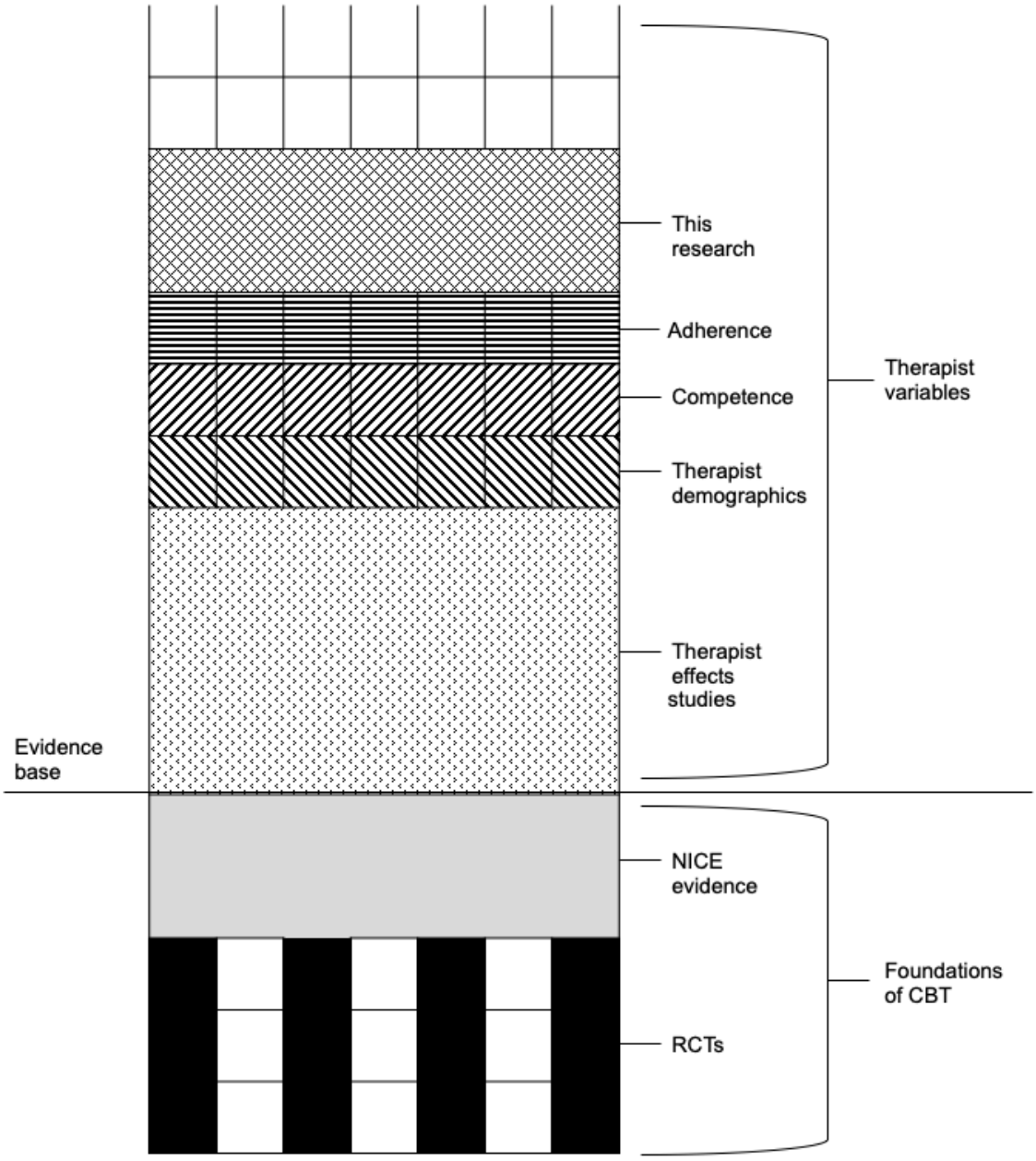
6.1 INTRODUCTION

This thesis began with an introduction to the problem that there is a significant variance in the reported recovery rates the United Kingdom's Improving Access to Psychological Therapy Programme. The objective of this research was to build on the current hypotheses about the reasons why there is such a variance in outcome. Chapter 3 of this thesis reported that, primarily, there are three types of research investigating the variables that may relate to outcome in CBT in general and also in High Intensity CBT therapists in IAPT. The first type of research, published in the literature, are secondary analyses of outcome trials (usually randomised controlled trials) whereby researchers investigate whether therapist adherence to a treatment protocol relates to outcome. The second type is process-outcome research (usually naturalistic observational cohort studies) whereby researchers are interested in the relationship between therapist variables and outcome. The third type of research is therapist effects studies where researchers investigate to what extent the combination of all therapist variables effects clinical outcomes. The research presented in this thesis is situated in the process-outcome research category and is a naturalistic observational cohort study. Figure 6.1 shows where this research is situated in relation to the research that has been conducted within IAPT with a particular focus on High Intensity cognitive behavioural. The foundations of IAPT are built on a strong evidence base that supports the efficacy of CBT (Clark, 2014). In the intervening years between the development of IAPT in 2009 and the current-time research has investigated the effectiveness of the IAPT programme. This has included research on clinical outcomes, therapist training and therapist demographics. Much of the previous research has reported on the variance in outcomes between services and between therapists, utilising the large data set reported to NHS England by IAPT services. One of the themes reported in this research is the significant variance in outcomes and this has led to a number of studies that investigate to what extent therapist effects explain the variance in outcome. Despite the growth in research within IAPT very little is known about

why some therapists achieve better outcomes than others and which therapist variables may be more significant than others. The findings reported in this thesis build on previous research and make a unique contribution to knowledge. This is the first research study to use therapy transcripts of complete episodes of care delivered by (n =200) High Intensity cognitive behavioural therapists. These transcripts have been used to investigate the therapist variables that relate to clinical outcome.

This chapter will discuss the findings from this research. The chapter proceeds with a discussion relating to where this research sits within the existing literature, particularly process outcome research and therapist effects research. This will be followed by a discussion addressing each research question and the limitations and strengths of this research. It will be argued that this research has made a significant contribution to knowledge in relation to how therapist competence and adherence to an evidence-based protocol are related to clinical outcome. Whilst the findings from this research have been presented as hypotheses in previous studies, the research reported in this thesis is unique in that this is the first time any study has reviewed the therapy transcripts from every therapy session of 600 patients who had all completed treatment at step 3 in IAPT. It will be argued that this research represents the emergence of a new way to conduct psychological therapy research. This research significantly adds to the literature as the first study to review therapy transcripts at this scale in IAPT. This chapter will argue that the findings from this research have the potential to have a significant impact on High Intensity CBT therapists, Clinical Supervisors, IAPT services, IAPT training programmes and NHS policy. The chapter will conclude with a discussion of the implications of the findings of this research on professional practice for each of these groups.

Figure 6.1 This research in the context of the research relating to IAPT. The figure is a metaphorical representation of how the relevant literature has been built on top of the evidence-base for CBT (horizontal line across the diagram). The evidence base represents the foundational structure which underpins research in the context of IAPT. Gradually new layers of research are added to the structure including, therapist effects studies, therapist demographics studies and the research reported in this thesis.



6.1.1 Research *in silico*: Studying therapy transcripts

The existing studies in the CBT and IAPT literature can be categorised in terms of the historical development of IAPT. The research, as shown in figure 6.1, is built on the foundations of a number of randomised controlled studies demonstrating the efficacy of CBT treatment protocols for a range of mental health disorders (Layard and Clark, 2014, McHugh and Barlow, 2012). These studies have informed the NICE guidelines for the treatment of common mental health disorders such as depression and anxiety disorders (Layard and Clark, 2014) and, as such, form the evidence-base that has become the curriculum for IAPT training programmes. Since the advent of the IAPT programme in 2009, researchers have reported on the effectiveness of the programme. This research falls broadly into three camps; those that largely lobby for the IAPT programme, those that have a more objective stance and those that are more critical of the programme. Many of these studies investigate the variance in outcome between services and between therapists. Unfortunately, these studies use different criteria for defining clinical outcome. This has made it difficult to compare the findings. Furthermore, some studies have repurposed the data from randomised controlled trials that had another research focus. It has been recommended that future studies are specifically designed as therapist variables or therapist effects studies (Baldwin and Immel, 2013). Secondly, many of the studies have been underpowered using far fewer than the recommended 100 therapists (Maas and Hox, 2004). Thirdly, and more importantly, all of the studies, to date, have had limited or no access to live therapy data. This is frequently cited as a significant limitation of the research in this field (Kuyken and Tsivrikos, 2009, Pereira, Barkham, Kellett and Saxon, 2017, Saxon, Firth and Barkham, 2017, Bruijniks, Franx and Huibers, 2018, Firth, Saxon, Stiles and Barkham, 2019). Even when researchers have access to large sample sizes (as in the Saxon, Firth and Barkham, 2017, analysis of 39.520 IAPT patients), without access to live therapy material it is not possible to understand the relationships between what therapists do and clinical outcomes. Therefore, analysing outcome data can only address part of the question of why some

therapists achieve higher clinical outcomes than others. Where researchers have had access to small numbers of recordings of therapy sessions, a further limitation is evident whereby researchers have allowed therapists to self-select which recordings they submit for review (see Branson, Shafran and Myles 2015). Self-selection of therapy recordings is potentially problematic as therapists may select recordings which they believe to be superior and these may be less representative of their day to day clinical work with patients (Mathieson, Barnfield and Beaumont, 2009).

The research reported in this thesis has considered the limitations of previous research in relation to lack of access to live therapy material (Bruijniks, Franx and Huibers, 2018) small sample size (Maas and Hox, 2004, Schiefele et al., 2016), research in real-world settings (Brosan, Reynolds and Moore, 2006) therapist self-selection of therapy sessions (Keen and Freston, 2008) and the number of raters used to assess competency and adherence (Brosan, Reynolds and Moore, 2006). Therapist adherence and competence are rarely investigated as potential predictors of outcome (Weck et al., 2016). This is the first study to have access to sufficient live therapy material, in the form of transcripts derived from IECBT sessions, to investigate whether competence or adherence correlate with clinical outcome. This new method of investigating therapist variables differs from the normal practice of studying or analysing therapy transcripts or therapy recordings on a computer. Research *in* silico, in the context of the research reported in this thesis, means that the therapy transcripts are created digitally, stored digitally and analysed digitally. This method of research is more normally used in the pharmaceutical and biological industries where research is conducted using digitally generated models rather the more common, *in vitro*, laboratory-based approach (Manning, Sleator and Walsh, 2012). Research *in silico*, in the pharmaceutical and biological industries involves computational models and very large data sets (Gunturi, Narayanan and Khandelwal, 2006) and is used in drug and treatment discovery (Walsh, Carroll and Sleater, 2013). In psychological therapy, research is more usually conducted *in vivo*, meaning that experiments and enquiry are undertaken on people

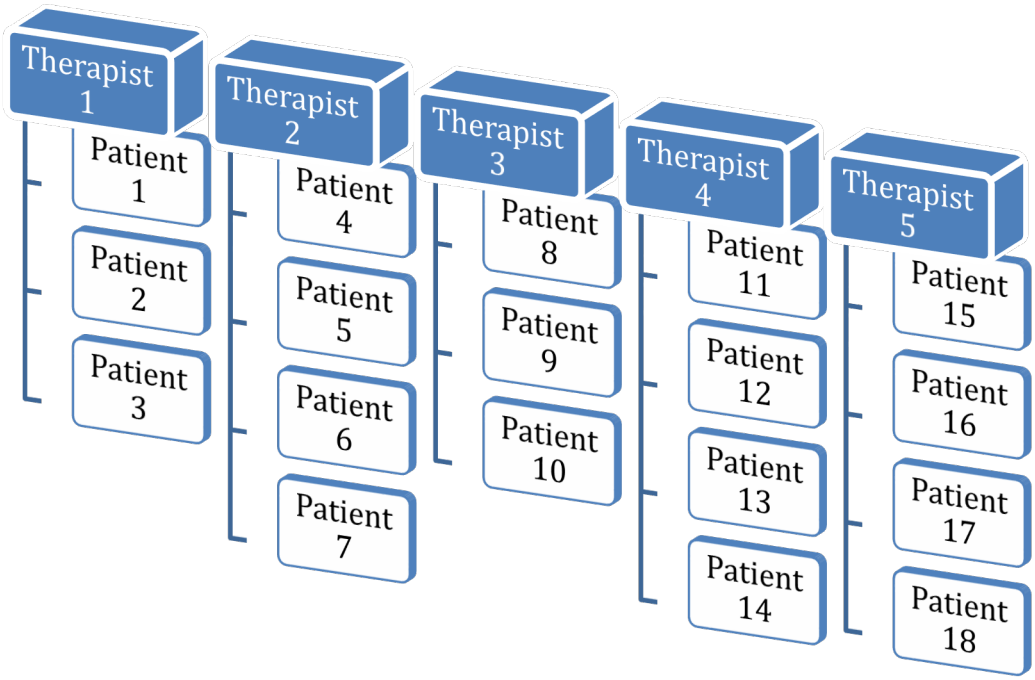
(patients and/or therapists). Research *in silico* represents a new method by which psychological therapy can be studied. This method makes it possible to study what therapists are doing, with every patient, in a real-world clinical setting and at scale. The only other field in psychological therapy that frequently uses large volumes of data are the therapist effects studies. However, this is unlike research *in silico* in that the data, which has not been created using a digital method, is transferred to a computer and then analysed. This is in contrast to the analysis of therapy transcripts that have been created digitally by over 600 BABCP accredited therapists since 2012. This digitally created data set can be used to study therapist (and patient) behaviour, enabling researchers to study, in large volume, real-world clinical data.

6.1.2 Therapist effects research using multi-level models

The IAPT programme has enabled researchers to analyse large volumes of outcome data (Layard and Clark, 2014). Whilst the volume of data has not included access to therapy recordings or transcript, the resulting research has led to an acknowledgement that therapist variables account for some of the variance in outcome reported by IAPT services. In the last 5 years, researchers that have investigated the variance in clinical outcomes in evidence-based psychological therapies have increasingly favoured multi-level model statistical analyses (Johns, Kellett, Saxon and Barkham 2019). Multi-level models (MLM), sometimes termed hierarchical level modelling (HLM) originated in the education sector where data is said to be grouped into hierarchical levels, such as students, class, school and area or town (Raudenbush and Bryk, 2002). MLM models are useful in analysing data where the independent variables are situated at different levels and the data is said to be ‘nested’ (Johns, Barkham, Kellett and Saxon 2019). For example, in an education setting, students are ‘nested’ within classes and classes and are ‘nested’ within schools, and schools are ‘nested’ within areas or towns. These hierarchical groups also exist in psychotherapy where,

at the lowest level, patients are ‘nested’ in therapists and therapists are ‘nested’ within services (see figure 6.2).

Figure 6.2 Patients nested in therapists



A growing number of studies have used MLM to explore data, that is nested in this way, in order to understand the comparative impact that therapists have on patients. These studies are known as therapist effects studies (Lutz and Barkham, 2015). At the time of writing there have been at least 118 therapist effects studies, the most significant of which are presented in two meta-analyses of therapist effects studies. These were undertaken by Baldwin and Imel (2013) and, more recently, Johns, Kellett, Saxon and Barkham (2019). Johns et al., report a therapist effect of between 0.2% and 29% with a weighted average of 5%, which is comparable to Baldwin and Imel’s (2013) earlier findings of between 3% and 7%. However, there is significant heterogeneity in the studies included in the meta-analyses making it difficult to disentangle the findings. Many of the studies were randomised controlled trials

where therapist effects were a secondary outcome. In addition, several of the studies were underpowered with a range of 6 – 1800 therapists where 13 of the studies had less than 100 therapists (Schiefele et al., 2016, Johns et al., 2018). However, all of the studies, except one (Goldsmith et al., 2015) reported a significant therapist effect. These findings merit further research in order to understand how to improve patient outcomes (Johns et al., 2018). Whilst this statistical method provides an indication of how much of variance between clinical outcomes is due to therapist effects, MLM does not explain which therapist variables are more significant than others. The literature review, contained in this thesis, indicated that whilst there is a growing body of work that relates to therapist effects in IAPT, very little is known about why some therapists achieve better outcomes than others. The scarcity of knowledge in this area is due to the practical problem of accessing sufficient recordings of live therapy sessions that can be analysed in order to understand what therapists may, or may not, be doing with their patients and how that might relate to clinical outcome. This identified gap in knowledge, and the growing knowledge relating to therapist effects, led to a decision to not include a therapist effects analysis in this research but instead to focus more specifically on therapist variables. This led to four research questions on which this thesis is based. The findings, relating to each of these questions regarding therapist demographics, therapist training, therapist competence and therapist adherence are summarised in the sections that follow.

6.2 RESEARCH QUESTION ONE: THERAPIST DEMOGRAPHICS & OUTCOME

The first research question, defined in Chapter 3 of this thesis, sought to understand whether a therapist's age, gender, years of experience or core profession correlates with clinical outcome. Initial research on therapist variables understandably focussed on these four variables (see Stein and Lambert 1984, 1995). It would seem a reasonable hypothesis that therapists with more life experience, more clinical experience or higher academic qualifications, such as a Doctorate in Clinical Psychology, would achieve better outcomes

than their younger, less-experienced, less-qualified colleagues. Whilst there was some early conjecture that supported this hypothesis, these studies had small sample sizes and tended to be based on trainee therapists rather than qualified therapists. One study reported that IAPT trainees who were Clinical Psychologists achieved higher marks for interpersonal effectiveness on the CTS-R (Brosan, Reynolds and Moore, 2006). It should be noted that interpersonal effectiveness is only one of twelve items on the CTS-R and the authors failed to correlate CTS-R scores with clinical outcomes. Despite some conjecture, it appears to be commonly agreed that therapist demographics such as age, gender, years of experience and training are not related to treatment outcome (Weck et al., 2015, Castonguay and Hill 2017). The findings reported in this thesis supports the existing literature in that a therapist's age, gender, years of experience or core profession were not found to be related to outcome.

6.3 RESEARCH QUESTION TWO: HI IAPT TRAINING & CLINICAL OUTCOME

The High Intensity cognitive behavioural therapist IAPT training programme (see Chapter 2) was developed to train a workforce that could deliver National Institute of Health and Social Care Excellence (NICE) recommended interventions with a significant focus on adherence to evidence based protocols (NHS England, 2018). High Intensity therapists working in IAPT may have either completed the IAPT training programme or may have completed equivalent training elsewhere (mainly prior to IAPT). Whilst there are studies that have investigated the effectiveness of the IAPT training programme (see Macmanus, Westbrook, Vazquez Montes, Fennel and Kennedy, 2010, Clark 2018, Liness, Beale, Lea, Byrne, Hirsch and Clark 2019) no studies, to date, have compared outcomes between therapists who completed the IAPT training and those that have not. It might be argued that the IAPT training programme was developed in order to enhance and improve the training for CBT therapists (Liness et al., 2019) and, therefore, IAPT trained therapists might be expected to achieve better outcomes.

The second research question in this thesis asks whether IAPT trained therapists achieve superior outcomes to non-IAPT trained therapists. This research found that there was no statistically significant correlation between therapists who had received IAPT training and clinical outcomes. In this sample of therapists, it would appear that patients had a similar chance of improvement whether they were seen by an IAPT trained therapist or not. One hypothesis for this finding is that therapists providing CBT online using IECBT are provided with access to an e-learning platform (see Chapter 1) which provides 'top-up' training based on the IAPT training curricula. This additional training may, in part, neutralise any difference between therapists who have undertaken IAPT training and those that had not, if, indeed, a difference existed in the first place. However, further research would be needed to analyse to what extent therapists, who had not undertaken IAPT training, were using the e-learning platform, compared to those that had completed an IAPT training programme.

6.4 RESEARCH QUESTION THREE: THERAPIST COMPETENCE & OUTCOME

Whilst it has been commonly assumed that therapist competence is related to clinical outcome (Trepka, Rees, Shapiro, Hardy and Barkham, 2004, Brosan, Reynolds and Moore, 2006, Strunk, Brotman, DeRubeis and Hollon, 2010, Branson, Shafran and Myles, 2015, Branson and Shafran, 2015, Braun Strunk, Sasso and Cooper, 2015) there is surprisingly little research that proves or disproves this assumption. The dearth of research in this area relates to the practical and logistical difficulties associated with accessing recordings of therapy sessions in real-world settings. Where research exists, there are significant limitations including repurposing the data from randomised controlled trials (see Weck, Richtberg, Jakob, Neng and Hofling, 2015, Ehlers, Hackmann, Grey, Wild, Liness, Albert, et al., 2014, Youn, Xiao, Kim, Castonguay, McAleavey, Newman and Safran, 2017), small sample sizes (see Brosan, Reynolds and Moore, 2006, Weck, Richtberg, Jakob, Neng and Hofling, 2015, Youn et al., 2017,) or therapist self-selection of therapy tapes. For example,

Liness et al., (2019) reports (in a cohort of IAPT trainees) that a 1-point increase in the CTS-R score relates to a 0.76% improvement in a patients PHQ-9 score at the end of treatment. The authors conclude that higher CTS-R scores predict improved outcomes on the PHQ-9. However, the study allowed trainees to self-select the therapy sessions that were rated on the CTS-R, thus reducing the likelihood that the findings were representative of the trainees overall practice (Ginzburg et al., 2012, Liness et al.,2019). Additionally, this study investigated trainees enrolled on an IAPT training programme rather than qualified therapists. Therefore, the recordings that the trainees were submitting for review were part of their formative and summative assessment and there was a requirement that the trainees needed to achieve a score of at least 50% in order to pass the course (Liness et al., 2019). The authors themselves acknowledge that this limits the generalisability of their research and, therefore, it remains unclear whether competence is associated with outcome in qualified therapists.

The third research question in this thesis asks whether therapist competence (as defined by the CTS-R) is related to clinical outcome. Using a simple Pearson's Correlation this research found that there was a statistically significant positive correlation (r= 0.167, p 0.019) between the mean of three randomly selected CTS-R scores (termed the F score in this thesis) and recovery. Whilst the answer to this question was affirmative, this research went on to use a linear regression model in order to investigate whether higher scores on individual CTS-R items predict outcome (see section 2.5 for further explanation of the individual CTS-R items). The model failed to predict a statistically significant effect of any of the individual CTS-R items on recovery. Therefore, it is presumed that, in this sample, none of the individual CTS-R items are more predictive of recovery than others, but the combined effect of all 12 items of the CTS-R is correlated to recovery. In order to further test the strength of this relationship a hierarchical multiple regression model was used whereby each of the variables were added to the model sequentially in order to test the strength of the F score (the mean of 3 CTS-Rs) when all of the other variables are added to the model. It was unsurprising that the

whole model (which included all the independent variables) failed to be statistically significant as it had already been ascertained that some of the variables (therapist demographics and therapist training) were not related to recovery. However, the best fitting model, which included just F scores and adherence scores (defined as A scores in this thesis) was significant at the 0.05 level (p= 0.03). The findings from each of these analyses suggest that CTS-R is positively correlated with recovery but at the 0.05 level.

A final statistical analysis was included in this research to further examine the associations between the variables. A hierarchical log-linear analysis was conducted to determine the associations between therapist F score and A score, patient variables and clinical outcomes. This analysis is discussed later in this chapter in section 6.2.

The affirmative findings of this research, that competence is related to outcome, supports a widely held assumption that competence (as rated by the CTS-R) is associated with clinical outcomes (Trepka, Rees, Shapiro, Hardy and Barkham, 2004, Brosan, Reynolds and Moore, 2006, Strunk, Brotman, DeRubeis and Hollon, 2010, Branson, Shafran and Myles, 2015, Branson and Shafran, 2015, Braun Strunk, Sasso and Cooper, 2015). This research addressed the limitations identified in previous studies by increasing the sample size, not allowing therapists to self-select therapy recordings and using data from a real-world setting rather than repurposing data used in a randomised control trial. The findings from this study are similar to Liness et al., (2018), who also report a relationship between competence and outcome at the 0.05 level. However, these studies did not use recovery as the independent variable, and it might be argued that other methods of measuring outcome such as reliable change (as in the Liness et al., study) are easier to achieve. Furthermore, it may not be meaningful to compare this research to other studies because previous research has either investigated competence in trainees, in therapists providing treatment as part of a randomised controlled trial, or in therapists who may not be delivering CBT.

A second factor that should be considered is the heterogenous nature of CBT. The CTS-R was designed to measure Beck's (1979) Cognitive Therapy for depression. However, cognitive behavioural therapists in IAPT may deliver other types of CBT such as Behavioural Activation (Jacobson, Martell and Dimidijian, 2001), Schema Therapy (Young, 1999), Compassion Focussed Therapy (Gilbert, 2010), Acceptance and Commitment Therapy (Hayes, 2004) and Mindfulness Based Cognitive Therapy (Teasdale, Segal, Williams, Ridgeway, Soulsby and Lau, 2000) as well as a range of disorder-specific protocols for anxiety disorders. The CTS-R may not be the most effective instrument to assess competence where a therapist is using a therapy other than a traditional Beckian approach. Whilst others have also mooted this as a possibility (Muse and McManus, 2013) there is currently no evidence that the CTS-R should not be used to assess competence across all methods of delivering CBT.

Thirdly, there is significant argument about the inter-rater reliability of instruments such as the CTS-R and its predecessor, the CTS (Schmidt, Strunk, DeRubeis, Conklin and Braun, 2018). Factors such as the rater training and experience in using the CTS-R, fatigue when rating or subjective biases may all influence how the rater scores each therapy session. Whilst the effort to reduce differences in training between raters has been described elsewhere in this thesis it must be emphasised that it is difficult to mediate for rater fatigue or bias and, therefore, the findings reported in this study should be interpreted with this in mind. Regardless of these potential difficulties defining therapist competence and understanding, the active ingredients of CBT is key to improving patient outcomes (Webb, DeRubeis and Barber, 2010) and the findings reported in this thesis regarding therapist competence provide a unique contribution to knowledge in this area.

6.5 RESEARCH QUESTION FOUR: THERAPIST ADHERENCE & OUTCOME

This thesis has argued that therapist adherence is more effectively assessed when a rater reviews every therapy session in a completed episode of care for at least three patients. Assessing therapist adherence involves a rater assessing whether or not a therapist has delivered therapy using an evidence-based treatment protocol, not just in one single therapy session but for the duration of the entire episode of care for that patient (see Chapter 2 for a full description of evidence-based protocols used in CBT). This thesis has explored how it is problematic to assess whole episodes of care because of the practical difficulty of accessing the data in real-world clinical settings. If it is difficult to access individual therapy sessions, then it follows that it is even more problematic to access live therapy material for every treatment session. Consequently, there are very few studies that examine the relationship between adherence and outcome, and those that do tend to be in the context of a randomised controlled trial where therapists are trained to adhere to a protocol in order that the efficacy of the intervention can be determined (Weck et al., 2016). Not surprisingly, many of these studies report that adherence relates to outcomes (see Ginzburg et al., 2012, Weck et al., 2016 and Haug et al., 2016). Whilst Gyanni, Shafran, Layard and Clark, 2013, Clark (2018) and Liness et al., (2019) report that therapist adherence relates to outcomes in IAPT their findings are based on the supposition that IAPT therapists are delivering NICE approved, evidence-based protocols. The authors fail to support their argument by extensively reviewing therapy sessions and admit that they do not have access to any recordings of therapy sessions. Therefore, it reasonable to state that there are no real-world studies that examine the relationship between adherence and outcome within IAPT. The fact that other than adherence/outcome studies reported in individual RCTs that there have been no previous attempts to assess adherence and as such there is limited criteria in relation to how adherence can be assessed in real-world settings. The researcher has used the best available guidance in relation to IAPT in the form of the evidence base used to inform the national curriculum and related summative assessment criteria within the IAPT training programme. However, it should be pointed out that trainees are not assessed on their ability

to adhere to an evidence-based protocol using recordings of every therapy session. Therefore, the method used to assess adherence described in this thesis has never been used before, and the resulting findings should be interpreted with this in mind.

The fourth and final question in this thesis asks whether therapist adherence (defined as the A score) to an evidence-based protocol is related to outcome. This research initially conducted a simple Chi Square test for association between the A score and recovery. There was a statistically significant relationship at the 0.05 level ($X^2 = 8.24$, $p= 0.04$). Whilst the answer to this research question was affirmative, this research then went on to further test the strength of the relationship between A scores and recovery. Using the same analyses as those conducted for F scores (see section 6.1.4, above), A scores were placed in a hierarchical multiple regression model with all the other variables. The model, which included both A and F scores, was significant at the 0.05 level. The findings from both these analyses suggests that adherence is related to outcome at the 0.05 level. This is the first time that therapy transcripts of every treatment session of 600 patients have been reviewed in order to investigate whether therapists are adhering to an evidence-based treatment protocol in IAPT and how adherence relates to outcome. The descriptive statistics in themselves are of great interest but should be interpreted with caution given that the method has never been used before. Figure 6. 3 shows the distribution of A scores across the therapist sample in this research.

Figure 6.3 Distribution of A scores, showing those therapists (in red) who demonstrated little or no adherence in the 3 episodes of care assessed

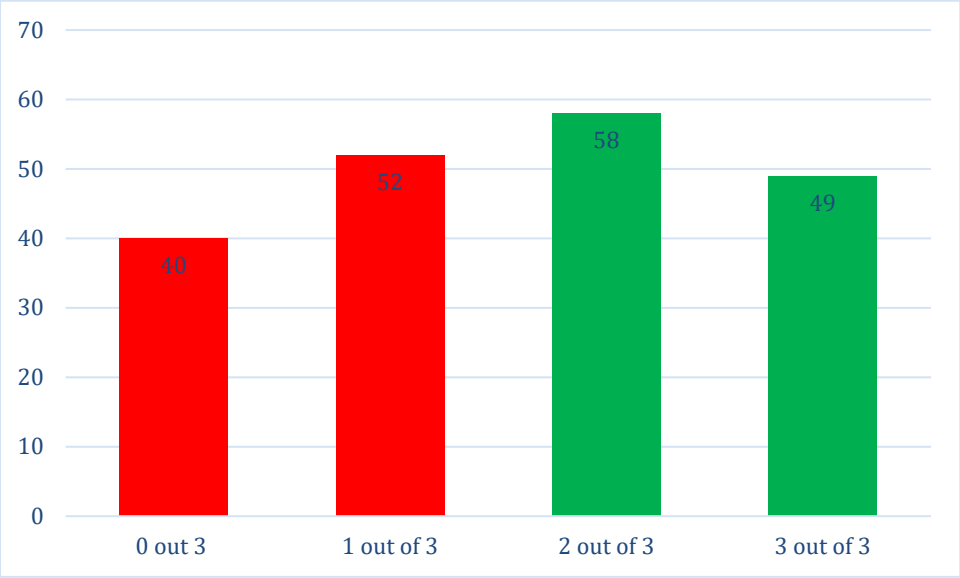


Figure 6.3 shows that 46.5% (n= 92) of the therapists in this sample were identified as demonstrating no adherence to a protocol in any of their treatment sessions or were adhering to a protocol in just one of the cases that were reviewed. This finding conflicts with that reported by Gyani et al., 2013, Clark 2018 and Liness et al, 2019 who suggest that CBT therapists closely adhere to an evidence-based protocol in IAPT. It should be noted that Liness et al., (2019) did not rate whole episodes of care and Gyani et al., (2013) and Clark, (2018) relied on therapists' self-reporting adherence to a protocol. The findings from this research reflect the arguments cited by Waller and Turner (2016) who suggest evidence-based psychological therapies are frequently delivered poorly. They suggest that therapists frequently drift away from delivering an evidence-based protocol due to a range of therapist characteristics, such as lack of knowledge, fear or a poor self-efficacy, whereby the therapist is concerned that they do not have the required skill set or attributes to follow the evidence base. Waller and Turner also suggest that therapists may believe that they are adhering to a protocol when they are not. It would appear that data derived from real-world research might differ from that reported in randomised controlled trials where therapist adherence is more rigidly controlled and monitored. If a significant minority of therapists are failing to adhere to

the evidence base and randomised controlled trials suggest that protocol adherence relates to patient recovery, then it might be argued that therapists in real-world settings need to encounter similar rigour (in terms of therapist selection, therapist training and therapist monitoring) in order to achieve similar outcomes. Alternatively, it might be argued evidence-based psychological therapy needs to move towards a new methodology, incorporating more real-world data in order that more patients can benefit from psychological therapies that make a meaningful difference to their lives.

The findings from the research discussed in this thesis are important. This is the first time that has been possible to examine the work of IAPT therapists in order to learn whether competence and adherence relate to outcome. The findings support the conjecture that many therapists fail to adhere to a protocol (Waller and Turner, 2016, Simmons, Milnes and Anderson, 2008, Cowdrey and Waller,2016, Stobie, Taylor, Quigley, Ewing and Salkovskis, 2007) and refutes the argument widely cited in the IAPT literature that therapists closely adhere to evidence-based protocols and NICE guidance (Gyani et al., 2013 and Clark, 2018). However, the findings from this research do appear to support the evidence base in that therapist adherence is correlated with clinical outcome at the 0.05 level. A final analysis, further testing the strength of this relationship, combining therapist A scores, F scores, clinical outcomes and patient variables is discussed in section 6.6, below.

6.6 DETERMINING THE ASSOCIATIONS BETWEEN THERAPIST VARIABLES, PATIENT VARIABLES AND OUTCOME

The findings from this research discussed in this thesis have so far determined that therapist demographics are not related to outcome but that both therapist competence (F scores) and therapist adherence to an evidenced based protocol (A scores) are related to recovery, albeit at the lower 0.05 level. In Chapters 2 and 3, this thesis explored some of the hypotheses that have been cited that explore the reasons for variance in reported outcomes

in IAPT. Whilst the focus of this research has been to examine what therapist variables may account for this variance, one other variable that is also likely to account for the variance are the patient variables. These include the patient's age, gender and severity of symptoms at the start of treatment (Catarino, Bateup, Tablan, Innes, Freer, Richards et al., 2018, Health and Social Care Information Centre, 2016).

The function of adding patient variables at this juncture is not to deviate away from the primary theme but to further examine the therapist A score and F score in the context of recovery and patient age, gender and severity. A hierarchical loglinear analysis was conducted as the final statistical analysis in this research. This additional analysis was unforeseen in the earlier stages of this research but was added in order to explore the differing effects between the variables. Hierarchical loglinear analyses require all variables to be categorical in nature (Everitt, 1977) and the data was converted to reflect this (see Chapter 5, section 5.6). The findings from the model are presented illustratively in figure 6.4. The model is divided into three sections to reflect the three categories of variables, where red indicates the therapist variables of A and F scores, blue represents the patient variables of age, gender and severity at the start of treatment on the PHQ-9 and GAD-7 (see Chapter 2 for descriptions of these outcome measures) and green represents the outcome variables of recovery and reliable improvement. Reliable improvement (as described in Chapter 2 of this thesis) was included in this model in order to examine whether this outcome variable had a stronger association with the predictor variables than recovery, which has been used throughout this research as the outcome variable. The illustration is used to highlight the strength of the association between the respective categories of variables through the numbers of lines connecting individual variables, where 3 lines represent a p value of $>$ 0.001. The model illustrates the strong association between therapist F scores and A scores. This is likely to suggest that therapists who demonstrate a higher level of fidelity to the model (F score) are also more likely to adhere to an evidence-based protocol, and vice versa. The loglinear model also confirms the previous findings, reported earlier in this

chapter, that fidelity to the model is associated with recovery. It is important to note that adherence (A score) was only associated with recovery via the F score (fidelity to the CBT model). This finding might be best understood in relation to the CBT training curricula (see Chapter 2) and the conceptual framework of CBT in general, whereby therapists are trained to structure sessions to incorporate the various elements that the CTS-R captures. These features are incorporated into therapy sessions regardless of the evidence-based protocol that the therapist has selected to treat their patients. It might be argued, therefore, that therapists need to focus on the delivering CBT with fidelity to the model, first and foremost, and the adherence to a protocol is secondary. That is not to say that it should be assumed that adherence is in some way inferior to a therapist fidelity to the model (F score) but that it is likely that a therapist needs to have high fidelity to the model in order to deliver an evidence-based protocol. This hypothesis is supported by earlier suggestions by Elkin (1999) and Kuyken and Tsivrikos, (2009), that competence is the channel through which other therapist variables effect outcome. As both A score and F score are closely associated, it would appear that both elements are important in their association with outcome, but that competence is a vehicle through which adherence is driven. Both variables require discreet skills but there is clearly some overlap between the two. The implications of this finding are discussed more extensively later in this chapter.

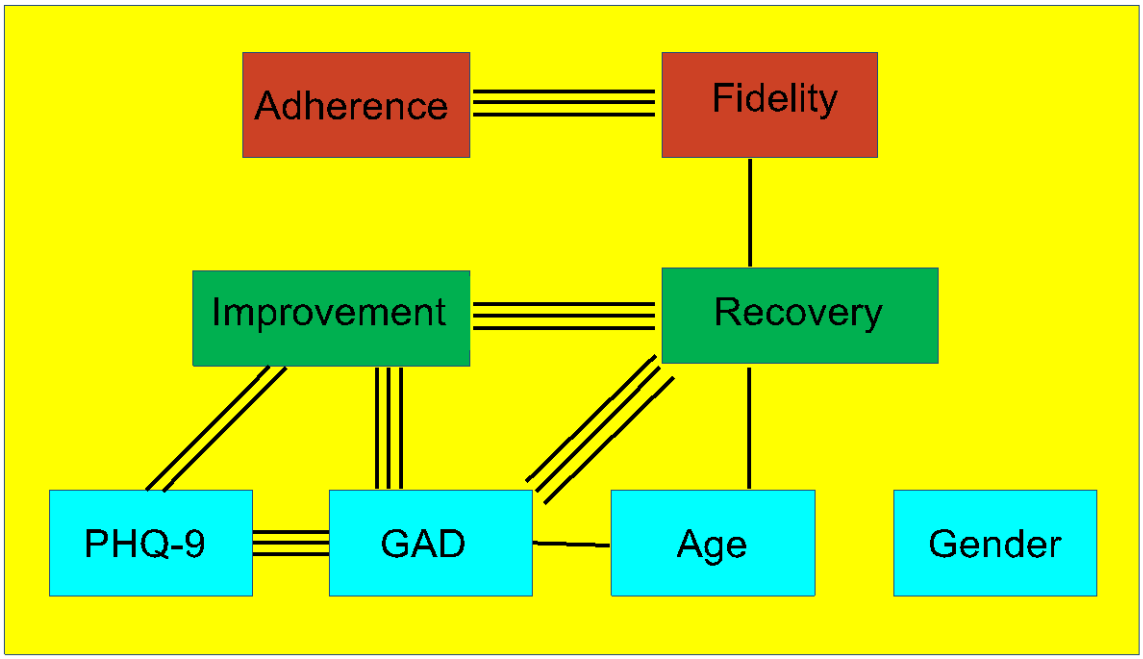
The findings relating to patient variables are less surprising and confirm what is already well-documented in the literature. The model illustrates that there is significant relationship between severity on both the PHQ-9 and GAD-7 at the < 0.001 level. That is to say, that patients are likely to present with similar levels of severity on both scales (Lamers, van Oppen, Comijs, Smit, Spinhoven, van Blakom et al. 2011) at the start of treatment. Severity, as measured by the GAD-7, is closed associated with recovery, also at the <0.001 level. This suggests that patients who present with an anxiety disorder (but potentially also with a similar score of severity on the PHQ-9) might be more likely to recover. This finding is

supported by Gyani, Shafran, Layard and Clark, (2013) who report higher recovery for patients who presented with Generalised Anxiety Disorder than those who presented with depression. Furthermore, the findings from the research reported in this thesis suggest that older patients (those over 60 years of age) are more likely to present with an anxiety disorder ($p = <0.01$) and older patients are more likely to recover ($p = <0.05$). These findings confirm those that are reported by Burch, Preston, Bateup and Hina (2018) and Catarino, Bateup, Tablan, Innes, Freer, Richards et al., 2018) and those that are reported by IAPT (NHS England, 2018). However, the current national data set reported by IAPT does not suggest there is a significant difference in recovery rates between patients who present with an anxiety disorder and those that are seeking treatment for depression (NHS England, 2018). Therefore, the findings in this thesis should be interpreted with caution, as they may only represent an artefact of the cohort of therapists/patients included in this research.

A final finding illustrated in the model is patient gender, which did not interact with any of the other variables. This might suggest that patient gender, on its own, is not a significant factor in predicting outcome. This finding is also reported by Cuijpers, Weitz, Twisk, Kuehner, Cristea, David et al., (2014) and suggests that gender is only important when it is considered in relation to other variables like age or severity. For example, it is widely reported that male patients under the age of 25 who present with greater severity are less likely to recover (Baker, 2018).

This final analyses have further confirmed the findings reported earlier in this thesis from the Pearson's Correlation, Chi Square Test for Association and regression models, whereby competence as measured by the CTS-R (F score) and therapist adherence to an evidence-based protocol are significantly related to clinical outcome at the 0.05 level.

Figure 6.4 Log-linear analysis model illustration



Therapist variables

Outcome variables

Patient Variables

p = <0.05

p = <0.01

p= <0.001

6.7 LIMITATIONS OF THIS RESEARCH

Various limitations of this research are identified, and these should be considered whilst interpreting the results and when planning future research.

6.7.1 Generalisability of the results

The IECBT method of delivering CBT is relatively new. At the time of writing, there are no other services, delivering this method of CBT. Therefore, it is not possible to benchmark the findings from this research against an IECBT comparative data set. While this thesis set out to demonstrate some similarities between traditional face-to-face CBT and IECBT (see discussion in Chapter 2, section 2.11) it is not clear whether therapists behave in similar ways online and in face-to-face settings. Further research is required to establish this. Therefore, whilst there is some assumption that the two methods of delivering CBT are broadly similar, both in terms of outcomes and the sample of therapists that largely represents those working in the national IAPT programme, results from this research cannot be assumed to generalise to other settings. In addition, as reported elsewhere in this thesis, it was difficult to reduce the likelihood of sample bias. 474 therapists, treating patients at Ieso Digital Health, were invited to participate in this study. Despite an attempt to recruit a large enough sample size, it is possible that only those therapists who felt confident about their clinical work agreed to participate. Therefore, it is possible that the sample is not wholly representative of the whole population. The potential for sample bias might have been reduced by waiting for more therapists to consent to participate in this research (thus increasing the sample size). Unfortunately, due to the time constraints of this research, this was not a viable option.

6.7.2 Data analysis

The data from this research was analysed quantitatively. Despite the focus on a well-defined and established outcome variable, type I errors cannot be excluded. The potential for type I errors might have been further reduced by increasing the sample size of therapists and increasing the acceptable number of patients treated by each therapist and therefore rating more sessions/completed cases per therapist. Increasing the number of sessions rated for each therapist may have assisted in the ability to more reliably assess therapist competency (Kazantis, Clayton, Cronin, Farchione, Limburg and Dobson 2018). However, undertaking these changes would have significantly extended the duration of this research and, consequently, it was not possible in the context of this professional doctorate.

Quantitative analyses were necessary in this research in order to compare and contrast the findings to existing research. However, additional qualitative analysis would have added another dimension to the findings. Given the availability of therapy transcripts, qualitative analysis might have included conversational analysis or thematic analysis of the transcripts in order to identify linguistic themes within the transcripts that might correlate with outcome. Whilst this was beyond the scope of this research it is a recommendation for future research and may yield important information that may add to the knowledge in this area.

6.7.3 Instruments used to assess competence and adherence

The CTS-R is the most widely used instrument to assess competence in IAPT (Liness, Beale, Lea, Hirsch and Clark, 2019). However, its ability to demonstrate whether or a therapist is competent is widely debated, particularly for qualified therapists (Trepka, Rees, Shapiro, Hardy, Barkham, 2004, Kazantis, Clayton, Cronin, Farchione, Limburg and Dobson, 2018). Competence in itself is something that is yet to be quantified, probably because of the difficulty in knowing what therapists are doing with their patients (Trepka et al, 2004, Hill and Castonguay, 2017). Therefore, it cannot be assumed that higher scores on the CTS-R relate

to competence in qualified therapists. Furthermore, there are no standardised or validated instruments available that assess therapist adherence to a disorder specific protocol. Primarily, this is because there are a number of treatment protocols, most of which are disorder specific. In the absence of an established instrument, this research used the Roth and Pilling (2007, 2008) competencies to form the basis of the rating of therapist adherence. The Roth and Pilling competencies form part of the IAPT training programme curriculum described in Chapter 2 of this thesis. It should be noted that IAPT training programmes do not use the Roth and Pilling competencies to assess adherence using live therapy material and, therefore, the use of this method is untested and subject to the opinion of the raters used in this research. Despite the fact that inter-rater reliability training was conducted, it cannot be assumed that each of the raters were able to be equally objective. Neither can it be assumed that this method of assessing adherence is appropriate or effective.

6.7.4 Inter-rater reliability

Whilst the intra-class correlation in this research was good, (ICC 0.98), it cannot be assumed that this eradicated subjective variance between raters. Whilst based on qualifying criteria, rating fidelity to the CBT model and adherence is likely to be subjective. Raters scores may be altered by many phenomena including tiredness, ill-health or the effect of rating one therapist's work after another, where the first therapist's work was poor which may make the subsequent therapist's work look superior (Schmidt, Strunk, DeRubeis, Conklin and Braun, 2018). Whilst it is not possible to eradicate this effect completely the effect might have been modified through the inclusion of rating each therapy session and treatment episode twice, using two separate raters, and then using a third rater to moderate the scores where there is a significant ($>$ 10%) difference between the scores. This process is similar to that used in the marking of IAPT trainees CTS-Rs (NHS England, 2018), however, it is time-consuming and was not possible in the scope of this doctoral research.

6.7.5 Outcome metrics

The PHQ-9 and GAD-7 were used as the primary outcome measures in this study, including the IAPT definition of recovery and reliable improvement. These outcome metrics and change indices were selected because their use is mandated within the IAPT programme and, therefore, it is possible to benchmark this research against prior and future research within IAPT. However, it is possible that these are not the most effective measures of outcome in that diagnosis specific measures, such as the PHQ-9, may not be the most reliable measure of progress for every patient. It has even been argued that the PHQ-9 and GAD-7 have not been rigorously tested in a primary care population of patients such as IAPT (Bohnke, Lutz and Delgadillo, 2014). Additionally, given that patients with depression may also have symptoms of anxiety and that features of depression vary from one patient to another (Krause, Lutz and Boehnke, 2011), then it may follow that other methods of measurement including transdiagnostic tools (Bohnke, Lutz and Delgadillo) or quality of life measures (Smits, Paap, and Bohnke, 2018) may be useful. It is evident that further research is required to ensure that services and researchers use tools that effectively measure patient outcome. Whilst this is beyond the scope of the research reported in this thesis, it is important to interpret the findings from this research with caution, bearing in mind that the PHQ-9 and GAD-7 may not be the most effective tools for measuring patient recovery.

6.7.6 Attempts to remove bias

Whilst it is not possible to completely eradicate bias, attempts to reduce it were incorporated into this research at each stage.

6.7.7 Selection bias

All therapists who had treated patients on the Ieso Digital Health platform were invited to participate in this research regardless of their clinical outcomes. Statistical guidance, in the

form of sample size tables, was drawn from Schiefele, Lutz, Barkham, Rubel, Bohnke, Delgadillo et al., (2016). Additionally, this research sought to exceed the sample size of all pre-existing process outcome research in the field of IAPT. Whilst this research followed the Schiefele et al., (2016) guidelines and exceeded the sample size of other similar studies in IAPT it might be argued that the sample (n=200) were self-selecting and therefore not wholly representative of the population.

6.7.8 Blinding

In an attempt to reduce the Hawthorne effect, whereby the therapists and/or the raters may be influenced by the purpose of a research study (Sedgwick, 2012), both raters and therapists were blind to the hypotheses for this research. In addition, to reduce the likelihood that the raters would be positively or negatively inclined towards the therapists' clinical sessions, the raters were blind to the clinical outcomes of each therapist's cases.

6.7.9 Observer bias

There is evidence that raters may have a positive bias towards a therapist they have previously supervised (Dennhag, Gibbons, Barber, Gallop and Crits-Christoph, 2012). To minimise such bias, raters were allocated therapists who were not known to them. Additionally, raters were asked to declare if the therapist was known to them and, where this was the case, the therapist was allocated to another rater. However, despite these attempts, it must be noted that was not possible to completely eradicate the possibility that the rater had some prior knowledge of the therapist and that this may have influenced the way they rated the therapists' work.

6.7.10 Measurement bias

This research used standardised validated instruments and outcome indices with the exception of the assessment of therapist adherence to an evidence-based protocol where no instrument is currently available. However, this research made attempts to base the assessment of adherence on quantifiable competencies that form the curriculum for the IAPT training programme and explicit and objective criteria were outlined to the raters during inter-rater reliability training.

6.8 STRENGTHS

This research has a number of strengths compared to others in the field. This is the largest naturalistic process-outcome study, to date, using live therapy material in addition to clinical outcomes and therapist demographics. This is the first time that transcripts of therapy sessions, for whole episodes of care, have been available in therapist variables/therapist effects studies in a real-world setting. Furthermore, this is the first time that it has been possible to rate adherence to an evidence-based protocol by rating the entire episode of care, rather than just one or two sessions. Additionally, this research is strengthened by the process of randomly selecting cases/sessions from a therapist's caseload rather than allowing a therapist to self-select a session for review. This research uses standardised validated instruments and outcome indices that are routinely used in all IAPT services, thus enabling the findings from this research to be used in the design of future research in this field. One of the problems cited that relates to existing research is the heterogenous nature of the studies, with a wide range of outcome measures, change indices, tools to assess therapist competence and statistical analyses being employed (Johns, Kellett, Saxon and Barkham, 2019). The data collected from this research has been analysed using a number of statistical methods, primarily focussing on those that have been commonly utilised in therapist effects/therapist variables research.

6.9 CONTRIBUTIONS TO KNOWLEDGE

This study sought to go beyond estimating how therapist effects account for the variance in clinical outcomes. Current research suggests that approximately 5-8% of the variability in clinical outcomes is due to therapist effects (Lutz et al., 2015). Whilst therapist effects research is a highly useful method to understand the variance in therapist outcome, it reveals very little about what the most, or least, effective therapists are doing with their patients. This limitation of therapist effects studies is a commonly cited phenomenon (Johns, Kellett, Saxon and Barkham (2019). Therefore, this study focussed on understanding whether therapist behaviour, in relation to fidelity to the CBT model and adherence to the evidence-base, is related to clinical outcome. This was only possible because of the availability of verbatim therapy transcripts for every therapy session delivered by each therapist who participated in the study.

6.9.1 Variance in outcome

It is widely accepted that therapists vary in their ability to effectively deliver interventions to patients, with some being far better than others (Lutz et al., 2015, Baldwin and Imel, 2013, Saxon and Barkham 2012). This research reports a similar finding with a significant variance between those therapists with the best outcomes and those with the worst outcomes. Figure 6.5 illustrates the variance between the therapists in this research, showing the rank position of each therapist.

Figure 6.5 Therapist rank position

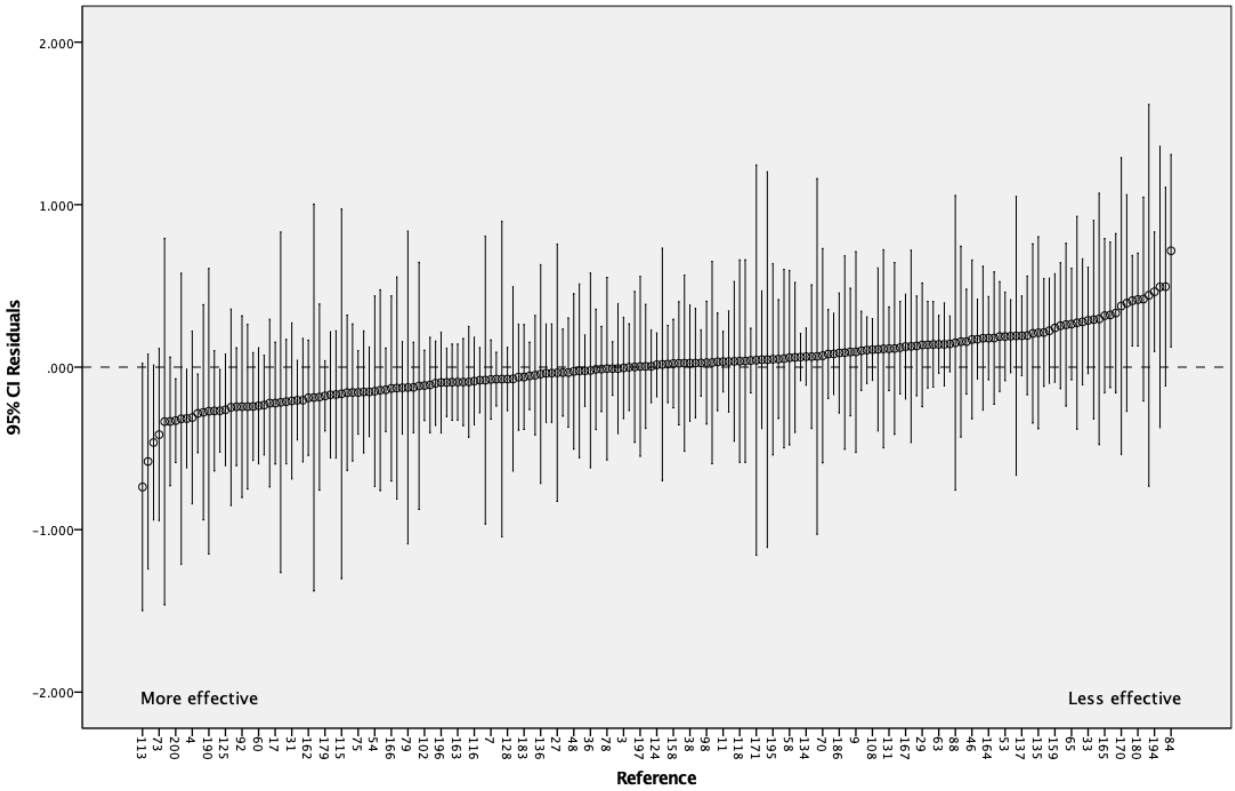


Figure 6.5 shows the recovery rate for each therapist in the sample ranging from the most effective to the least effective. The mean recovery rate for all the therapists was 50% (SD 19.02, SE 1.35) with a range of 0% to 100%. Understanding what the most effective therapists are doing with their patients has the potential to significantly impact on therapist training, continuing professional development and clinical supervision and may result in more patients reaching clinical recovery (Castonguay, Eubanks, Goldfried, Muran and Lutz, 2015).

6.9.2 Therapist Demographics

This research found that therapists demographics did not account for the variance in outcomes. This finding is unsurprising in that it is already well accepted that therapist demographics are unimportant in relation to outcome (Hill and Castonguay, 2017). Two noteworthy exceptions to these finding exist. Branson and Shafran, 2015 and Brosan, Reynolds and Moore, 2006 both report that those therapists with higher academic qualifications tended to score better on the CTS-R and academic writing (Branson and Shafran). However, neither study use clinical outcomes and therefore it is impossible to state whether higher academic qualifications would have corelated with outcome in their respective studies.

6.9.3 Therapist competence

This thesis has reported that therapist competence is associated with outcome and this supports a commonly-held hypothesis in the associated literature (see Chapters 2 and 4). The findings from earlier research are mixed with some studies suggesting that there is a relationship between competence and clinical outcomes (Liness et al., 2019, Haug et al., 2016, Brown et al., 2013, Strunk et al., 2010, Kuyken and Tsivrikos, 2009 and Trepka et al., 2004) and others refuting this (Branson, Shafran and Myles, 2015). It is important to note that the Branson et al., study used very minimal reliable change indices (a reduction of 2.12

on the PHQ-9 and 2.48 on the GAD-7) to report improvement. This is significantly different to the change indices used in this research where a patient had to fall below caseness on both PHQ-9 and GAD-7 in order to be defined as recovered (see section 2.12 for more details of the definition of recovery). Therefore, it is surprising that Branson et al., did not report a correlation between competence and outcome given that it was arguably easier to meet the reliable change criteria. That being said, the setting for the Branson et al., study was an IAPT training programme and, therefore, the therapists were novices rather than qualified therapists. It may not be possible to generalise the results from this study (and other research involving trainees) to qualified therapists (Branson, Shafran and Myles 2015, Liness, Beale, Lea, Byrne, Hirsch and Clark, 2019). The problem that has been identified in this research is that there are insufficient numbers of studies in this field that examine qualified therapist effects/variables, in real-world clinical settings, that have adequate access to live therapy material. Those that do, are so heterogeneous that it is difficult to draw conclusions. This research has sought to use outcome measures and processes that are common and easily replicable in order to add to the knowledge in this field. This research found that therapist competence was positively related to clinical outcomes at the < 0.05 level. This finding is the same as the Liness et al., (2019), Haug et al., (2016), Brown et al., (2013), Strunk et al., (2010), Kuyken and Tsivrikos, (2009) and Trepka et al., (2004) studies. However, only Liness et al., (2019) used PHQ-9 and GAD-7 (as outcome measures) and CTS-R to rate competence. Each of the other studies used different measures and different methods to assess competence. Despite these differences, this research confirms the findings of other studies whilst increasing the sample size and access to a random selection of therapy sessions.

6.9.4 Therapist adherence

This thesis reported that adherence correlated with outcome at the < 0.05 level. Whilst not causal, the findings in this thesis appear to support the assertions from prior research that

therapist adherence to an evidence-based protocol is significantly related to achieving good clinical outcomes (Gyani, Shafran, Layard and Clark, 2013, Ginzburg et al., 2012 and Shafran et al.,2009). This is a necessary first step towards establishing causality. However, this thesis has argued that caution should be taken when interpreting the findings from therapist effects/therapist variables research which has been conducted as part of research trial to determine the effectiveness of a particular treatment protocol. Randomised controlled studies tend to select therapists who are thought to be more likely to adhere to the protocol and therapists in these trials receive specific training and supervision in protocol adherence (see Ginzburg et al., 2012, Haug et al., 2016 Saxon, Firth and Barkham, 2017). Additionally, therapists delivering interventions in treatment efficacy or effectiveness trials are arguably more likely to make every attempt to adhere to the protocol (Roth, Pilling and Turner, 2010). Therefore, it is unremarkable when researchers report high levels of protocol adherence in randomised controlled trials and subsequently draw the conclusion that adherence correlates with outcome. That is not to say that researchers are incorrect in their assertions but that these findings are less likely to be generalisable in real-world settings (Roth, Pilling and Turner, 2010). Other studies have either not focussed on adherence or have rated either single sessions or small (< 3) numbers of sessions. One exception to this is Brown, Craske, Glenn, Stein, Sullivan, Sherbourne, et al., (2013) where raters had access to four sessions which had been randomly selected from each patient treated. Arguably, rating more than 2 sessions increases the likelihood that the rater will get a more accurate impression of a therapist (Trepka, Rees, Shapiro, Hardy and Barkham, 2004) as it is highly probable that therapist competence and/or adherence will vary from session to session (Minonne, 2008). The approach that Brown et al., (2013) have taken to research in this field is unusual and may only have been possible because the of the relatively small sample size (n= 14). Clearly it is more costly and time consuming to assess adherence from sessions selected from an entire course of treatment for a larger sample. It should be noted that the Brown et al., study is from the United States of America (USA) and is based on novice therapists so the results may not be generalisable to the UK IAPT programme. In addition,

Brown et al., developed their own instrument to rate therapist competence and adherence and therefore it is difficult to know whether their findings would have been the same had they used the CTS-R and Roth and Pilling Competencies to rate competence and adherence. Brown et al., report that that whilst competence was related to outcome at the < 0.05 level, adherence did not have a statistically significant relationship with outcome. It is important to note that CBT in the USA is arguably different to CBT in the United Kingdom's IAPT programme. At the time of writing, the USA does not have a national CBT curriculum, established minimum training standards or an accreditation process (Kobak, Wolitzky-Taylor, Craske and Rose, 2017). Most therapists have very little exposure to CBT training (Institute of Medicine, 2015) in the USA. More importantly the Brown et al., study repurposes data from a previous randomised controlled trial using computer-assisted, guided self-help (see Chapter 3 for a description of guided self-help) materials for anxiety. Therefore, it is difficult to draw many comparisons between The Brown et al., study and this research other than the finding, in relation to therapist competence, is similar. The research reported in this thesis builds on the Brown et al., study and extends the work of a further 3 studies (Webb, DeRubeis, Dimidjian, Hollon, Amsterdam and Shelton, 2012, Ginzburg, Bohn, Hofling, Weck, Clark and Stangier, 2012 and Haug, Nordgreen, Ost, Tangen, Kvale,Hovland et al., 2016) which examine the relationship between adherence and outcome. All three studies repurpose data from previous trials and, therefore, the results are arguably less generalisable. Webb et al., use data from two trials. The first trial has data from n= 6 therapists and the second has n = 3 therapists. The Ginzburg et al., trial had n= 10 therapists and Haug et al., had n= 22 therapists. All three trials used different outcome metrics and tools to assess adherence. However, Ginzburg et al., used a similar technique to the techniques described in this thesis to assess therapist competence in that they asked raters to decide whether the therapist had adhered to the treatment protocol, as described in the treatment manual for Cognitive Therapy for Social Anxiety Disorder. Raters in this trial were asked to review two tapes from each of 34 patients. Similarly, the other two trials asked raters to review only two recordings of therapy sessions, one drawn from the early stages of

therapy and from a session towards the end of therapy. It is unclear why a decision was made not to review more sessions in order to assess therapist adherence across an entire episode of care. Webb et al., and Haug et al., report that adherence was related to better outcomes. Webb et al., qualify 'better outcomes' by reporting that where the therapist adhered to Cognitive Therapy (it is not clear what protocol was being followed) that the patient acquired more skills. Arguably, this is very different interpretation of outcome. Conversely Ginzburg et al., report that adherence was not related to outcomes. Whilst all 3 studies provide useful and necessary insights into the relationship between adherence and outcome, the limitations of their studies brings into question the reliability of their findings as it impossible to understand whether the two sessions that were rated represented what occurred in the other sessions (Weck, 2014).

Given the lack of research in this area, the absence of research in an IAPT setting and the limitations outlined in the research that has been conducted the research presented in this thesis has made a significant contribution to the literature. In this research, not only were the therapists treating patients in a real-world clinical setting, but therapy sessions were selected randomly by the raters and the whole episode of care was assessed.

6.9.5 Relationship between competence and adherence

A further noteworthy finding of this study is the relationship between competence (fidelity to the CBT model as rated by the CTS-R and therapist adherence to an evidence-based protocol. Little is known about therapist behaviour in real-world settings (Brosnan, Reynolds and Moore, 2006). This is the first study to have access to the therapy transcripts of every patient treated by the therapists working in a clinical setting. Given that there is extensive literature that supports the idea that there is significant variance in therapist competence (Johns, Kellett, Saxon and Barkham, 2019) it is not surprising that the findings from this study also suggest this. Even studies that have not been conducted in real-world setting report this variance. Strunk et al., (2010) using data drawn from a randomised trial report a

variance in fidelity to the model with a range of 17.8% - 56.6% and mean score of 39.7%. Similarly, the variance in therapist adherence to a protocol is not uncommon. There are many possible reasons why therapists drift away from a protocol. These include negative beliefs about aspects of delivering treatment such as exposure (Deacon et al., 2013), which lead to avoidance of these aspects of therapy. Additionally, therapists make judgements when deciding which treatment methods to use and this may lead to a dilution of the protocol (Grove, Zald, Lebow, Snitz and Nelson, 2000) and inflated self-beliefs about competence (Parker and Waller, 2015). This, in turn, may negatively bias therapists from seeking effective continuing professional development (Parker and Waller, 2015). Whilst the phenomenon of therapist drift is widely documented it is important to note that previous research has not been based on the review and analysis of large volumes of live therapy data. Therefore, the research reported in this thesis provides significant evidence to support the hypothesis that therapist drift is a common phenomenon. Indeed, in their meta-analysis Zarafontis-Muller, Kuhr and Bechtdolf (2014) report a correlation between clinical outcomes and therapist competence ($r= 0.24$) and adherence ($r = 0.06$). More recently, Kaznatis et al., (2018) also reported similar findings. This study also finds that competence (as assessed by the CTS-R) is correlated with clinical recovery ($p<0.05$). An important finding is that adherence (A score) is related to clinical recovery through fidelity to the CBT model (F score). This might suggest that, first and foremost, delivering therapy that contains the core elements of CBT i.e. agenda setting, giving and eliciting feedback, guided discovery, conceptual integration etc, is highly important, but when therapists achieve this whilst adhering to a protocol, clinical outcomes improve. Metaphorically, fidelity to the CBT model might be viewed as the vehicle that is required to effectively deliver a disorder specific protocol and that, without the vehicle, the protocol alone cannot be delivered. If fidelity to the CBT model and adherence to a protocol are correlated with clinical outcome (albeit at the lower 0.05 level), then further work may be required to support therapists to amplify their work beyond what is currently available. On the other hand, it is important to acknowledge that whilst fidelity to the CBT model and adherence appear to be related to outcome it is

possible that the tools that are currently used to assess competence and adherence are unfit for purpose. Whilst the most contemporary evidence suggests that therapists should deliver CBT with fidelity to the model whilst adhering to an evidence-based protocol, there has been very little advancement in the evidence base towards new methods of delivering therapy that might produce better clinical outcomes. Despite the fact that, in the last 70 years, researchers have demonstrated that there is an evidence-base for CBT (McHugh and Barlow, 2012) very little is actually known about the active ingredients of CBT and why or how it works (Lorenzo-Luaces, German and DeRubeis, 2015). Therefore, it might be argued that until the mediators and moderators of CBT are well-established then it is not possible to build tools that can effectively rate therapist competence or adherence. Real-world research may expose the work of the best performing therapists so that it becomes possible to learn from these therapists and potentially advance the evidence base (Bruijniks, Franx and Huibers, 2018, Hill and Castonguay, 2017). Some early research in this field points to a new type of therapist drift which has more positive connotations (Bruijniks, Franx and Huibers, 2018). The authors suggest that some therapists adapt treatment protocols, flexing the protocol to meet the idiosyncratic needs of the patient in a way that does not sub-optimise treatment outcomes. Bruijniks et al., argue that adapting protocols to meet the needs of individual patients may be desirable and unlike therapist drift, as described by Waller and Turner (2016), this does not have a negative impact on treatment outcomes. Understanding how therapist adherence to evidence-based protocols impacts on patient outcomes is vital (Bruijniks et al., 2018, Waller and Turner, 2016).

6.10 THE NOTION OF FLEXIBLE ADHERENCE

Flexible adherence to an evidence-based protocol has been defined as the idiosyncratic adaptation of a treatment protocol so that it meets the identified needs of an individual patient (Bruijniks et al., 2018). In practice, this may mean that a therapist omits part of a protocol because it is not required, or part of a protocol is emphasised because it plays a

more significant role in the successful reduction of the patient's symptoms. Whilst the notion of flexible adherence is not fully understood, it would appear that flexible adherence is based on a clinical rationale rather than the drift that Waller and Parker (2016) describe. Furthermore, the highly skilled adaptations of a protocol used in flexible adherence may describe the work of the most highly effective therapists. The work of these most highly effective therapists has not been studied in vivo. Current research is based on patient (Ricks, 1974) or therapist reports (Bruijniks et al., 2018) of what the best therapists might be doing with their patients. It will be important to review the work of the most effective therapists in order to learn what they are doing. For example, if the most effective therapists are using flexible adherence then this has implications for the future of evidence-based psychological therapy.

In this thesis raters were asked to assess whether therapists had adhered to a protocol by reviewing all the treatment sessions of a patient's episode of care. Each of the raters had experience of teaching on IAPT training programmes. Given that adherence in the context of IAPT relates to the delivery of treatment following a specific evidence-base or set of competencies (for example Roth and Pilling 2007, 2008) then it might be argued that the raters were looking for rigid adherence and that those therapists who exhibited flexible adherence were marked down because not all elements of a protocol were evident. However, given the term 'flexible adherence' is, as yet, ill-defined it is difficult to assert this hypothesis with any confidence. Closer, qualitative, analysis of the transcripts of the most effective therapists may provide a better understanding of what these therapists are doing with their patients.

6.11 RECOMMENDATIONS FOR FUTURE RESEARCH

This research has reported that both fidelity to the CBT model, and adherence to an evidence-based protocol, are related to clinical outcomes. Huppert et al., (2001) raise the

question that if it were possible to understand the difference between effective and ineffective therapists that the next step would be to understand whether or not is possible to make ineffective therapists more effective.

1. This thesis has reported that there was significant variance between therapists, with some therapists achieving better outcomes than others. This finding is congruent with other studies. It is recommended that qualitative examination of the therapy transcripts of the best performing and worst performing therapists is undertaken in order to understand what the best/worst therapists are doing with their patients. Learning from this research might then be used to inform therapist training with the aim of helping poor performing or average therapists become more effective.

2. This thesis has reported that one limitation of this research is that it was not possible to assess more transcripts/whole completed cases. A recommendation for future research is to understand how many transcripts/whole cases need to be rated in order to be confident in the findings. Furthermore, given the debate relating to how to define and/or assess processes such as therapist competence and adherence further research might explore new tools and instruments that reliably report therapist activity.

3. This thesis has discussed how IAPT training programmes currently assess therapist competence and adherence. Further research might involve assessing competence and adherence of trainee therapists enrolled on an IAPT training programme using the IECBT method whereby academic tutors, assessor and supervisors could assess whether a trainee could consistently deliver CBT with fidelity to the model and adherence to a disorder specific treatment protocol. Research might test the feasibility of using the IECBT method as a formative assessment tool in order to

support academic staff, assessors and supervisors to highlight therapists who are struggling with key techniques. A further hypothesis, that this method might lead to improved outcomes and improved scores at summative assessment might then be tested.

4. This thesis analysed therapists who delivered CBT online using written communication. It is not yet possible to understand whether a therapist's behaviour online differs from their behaviour when delivering treatment face-to-face. Whilst IECBT and face-to-face CBT are comparable in terms of clinical outcomes and variance in therapist effectiveness, further research is required to understand any differences between the two methods. It is recommended that transcribed face-to-face CBT sessions and IECBT transcripts (delivered by the same therapist) are analysed in order to ascertain whether there are significant differences between the two methods.

6.11.1 Summary

This section has highlighted areas for further research based on the findings of this current research. The findings make a significant contribution to the body of work already conducted in this field, especially because some of the limitations cited in previous studies have been addressed. The delivery of evidence-based psychological interventions via Internet Enabled CBT (IECBT) has provided a new, and currently unique, way to learn more about how therapist behaviour impacts on good clinical outcomes. It has been argued that the methods used in the research reported in this thesis represent a new method of conducting research in psychological therapy. The implications of the findings from this research on routine clinical practice, the education and training of therapists and the provision of clinical supervision are discussed in the remaining sections of this chapter.

6.12 IMPLICATIONS FOR PROFESSIONAL PRACTICE

This thesis explored three main themes; therapist demographics, therapist adherence to an evidence-based protocol and therapist competence as rated by the CTS-R (fidelity to the CBT model). These themes have been explored in terms of their relationship to clinical outcomes in the context of the United Kingdom's Improving Access to Psychological Therapy programme (IAPT). Chapter 5 of this thesis discussed the findings from this research and established that both competence and adherence are related to clinical outcome at the <0.05 level, but therapist demographics (age, gender, years of experience, core profession and training) are not. The remaining sections of this chapter will discuss the implications of these findings on professional practice both in clinical settings and in the education and training of therapists.

6.12.1 Education and training of cognitive behavioural therapists

Chapter 2 of this thesis outlined the current curricula for IAPT High Intensity Therapy training programmes in the UK. Whilst the success of the IAPT training programme is not in doubt, it might be argued that there could be increased focus on the significance of the relationship between clinical outcomes and therapist adherence to the cognitive behavioural model and adherence to a protocol. Arguably, if trainee therapists were required to use Internet Enabled CBT (IECBT) to treat a proportion of their patients, the transcripts derived from these online sessions could be used to enhance trainee self-reflection and supervisory feedback to guide both the trainee and the training programme tutors on the trainees' progress towards delivering treatment in accordance with the IAPT curriculum. However, it should be emphasised that the hypothesis that IECBT transcripts would enhance learning and, therefore, improve therapist competence/adherence has yet to be tested.

6.12.2 The formative and summative assessment of competence

This research has established the feasibility of undertaking CTS-R ratings using therapy sessions that have been delivered via the Ieso Method (IECBT). The benefits of rating CTS-Rs using the IECBT are a reduction in the time it takes to rate one session by at least 50% (Ewbank, Cummins, Tablan, Bateup, Catarino, Martin and Blackwell, 2019) and the availability of every therapy session so that the rater can randomly select a session to review. Currently, IAPT trainees are required to submit recordings of 3 therapy sessions, each of which must be derived from a separate patient. Trainees self-select recordings to submit for assessment (Branson, Shafran and Myles, 2015) and it is unlikely that a trainee will submit a recording that they believe would not achieve the pass mark (50%). Therefore, the quality of the rest of their work remains unknown. Additionally, due to the small number of therapy sessions that are CTS-R rated the trainee is receiving very little formative feedback. Given the argument that the CTS-R is a tool that supports learning and, therefore, the development of clinical skills (Brosan, Reynolds and Moore, 2008) it would seem prudent to use the CTS-R more frequently as a formative assessment tool. However, rating a CTS-R and providing high-quality feedback to a trainee, requires a considerable resource (Keen and Freeston, 2008) and this would incur additional expense and resource. Incorporating the use of Internet Enabled CBT (IECBT) into the IAPT curriculum, whereby trainees were required to treat a percentage of their patients using IECBT, might mediate for this additional resource. If IAPT programmes were to introduce this method then it would be possible for programme tutors and supervisors to randomly select sessions for formative feedback, thus eliminating the trainees' self-selection bias. The additional benefits would be that an increased number of CTS-Rs could be undertaken and the trainee would also benefit from the ability to self-reflect on each therapy session that had been delivered via the Ieso Method. Furthermore, if IAPT programmes were to use the Ieso Method as a formative assessment tool to enhance competence, then a next step might be to use the method as part of the trainees' final summative assessment. It could be argued that adding a randomly

selected session (drawn from the trainees' IECBT caseload) would strengthen the summative assessment without adding significant burden to the programme's teaching team. This aligns with a general desire, in the education for health care professionals, to explore how to increase the sample of clinical cases that are assessed (Miller, 1990, Govaerts, van der Vleuten, Schuwirth and Muijitjens, 2007, Brown and Doshi, 2006, Keen and Freeston, 2008). The ideal and fundamental aim would be to develop a method whereby it was possible to continuously assess what health care trainees were doing with their patients in routine practice. Clearly the IECBT method of delivering CBT would move the IAPT training programme closer to that aim.

Regardless of whether IAPT training programmes choose to incorporate the Ieso Method into the formative and summative assessment process, the implications of the findings from this research remain important. Understanding that competence is related to outcome strengthens the theory that underpins the IAPT curriculum. This in itself may have an impact on curriculum design and the pedagogical processes used to train the therapists of the future.

6.12.3 The formative and summative assessment of adherence

Currently, IAPT training programmes assess a therapist's ability to deliver therapy with adherence to an evidence-based protocol through academic writing (Liness, Lea, Nestler, Parker and Clark, 2016). Trainees are required to submit case reports outlining the assessment and treatment processes of two patients that they have treated. Whilst the trainee is required to also submit one recording of a treatment session from this case, which is rated via the CTS-R, currently there is no clinical assessment of the therapist's ability to adhere to a protocol. Given that the findings from this research was that adherence to a protocol (A score) is highly associated with competence (F score) at the <0.001 level, and competence is associated with better clinical outcomes then it would seem relevant that one

implication of this research is that IAPT training programmes consider developing further ways of assessing adherence. However, this undertaking is not insignificant. IAPT programmes teach at least 10 separate, disorder specific protocols and assessing a trainee's ability to adhere to each one is a significant undertaking and one that would come at a cost, both in time and resource. This thesis has argued that it is necessary to review every session of a completed episode of care in order to assess whether or not a therapist has adhered to a protocol. Arguably, additional reviews of trainee's sessions would cause a burden to IAPT training programmes.

6.12.4 Improving outcomes for therapy delivered by qualified therapists

This research has reported that therapist competence (as rated by the CTS-R) is related to recovery at the <0.05 level and that therapist adherence to an evidence-based protocol is closely related to competence at the <0.001 level. These findings suggest that the best therapists are delivering CBT with fidelity to the model, whilst adhering to a protocol. This is a relevant and important finding because, in order to improve clinical outcomes and reduce the variance between therapists it will be important to understand what the most effective therapists are doing with their patients (Brown, Lambert, Jones and Minami, 2006).

6.12.5 Therapist competence and continuing professional development

There are relatively few tools that are routinely used to assess therapist competence in cognitive behavioural therapy and there are even fewer in other psychotherapy modalities, such as psychodynamic psychotherapy (Schmidt, Strunk, DeRubeis, Conkin and Braun, 2018). The majority of research in the field of psychological therapy is outcomes-based research which focuses primarily on demonstrating that one form of therapy is more effective than another type of therapy (Okiishi, Lambert, Nielsen and Ogles, 2003, DeRubeis, Gefland, German, Fournier and Forand, 2014). However, there is very little research which focuses on individual therapist competence and, as discussed earlier in this thesis, the

methodology and heterogenous nature of the studies make it difficult to draw significant conclusions (Schmidt et al., 2018). The majority of research in the competence-outcome field use either the Revised Cognitive Therapy Scale (CTS-R) or the Cognitive Therapy Scale (CTS). Both the CTS-R and CTS are validated instruments that are used in the UK and USA, respectively. It is recommended that High Intensity CBT therapists working in IAPT use the CTS-R for continuing professional development either by self-rating therapy sessions or submitting recordings of therapy sessions to their supervisor for CTS-R (Liness, Lea, Nestler, Parker and Clark, 2016). In practice very few therapists do this on a regular basis (Liness et al., 2016). It might be argued, therefore, that therapists are less likely to focus on the competencies that the CTS-R assesses. If a higher score on the CTS-R relates to outcome, then one implication of the findings from this research is on therapists' continuing professional development. Based on the findings from this research, cognitive behavioural therapists, clinical supervisors and, indeed, IAPT services might place more significant emphasis on the regular review of therapy sessions using the CTS-R. Supervisor feedback from the CTS-R and therapist self-reflection on action (Bennett-Levy, 2006) are thought to enhance therapist competence (Bennett-Levy and Lee, 2012).

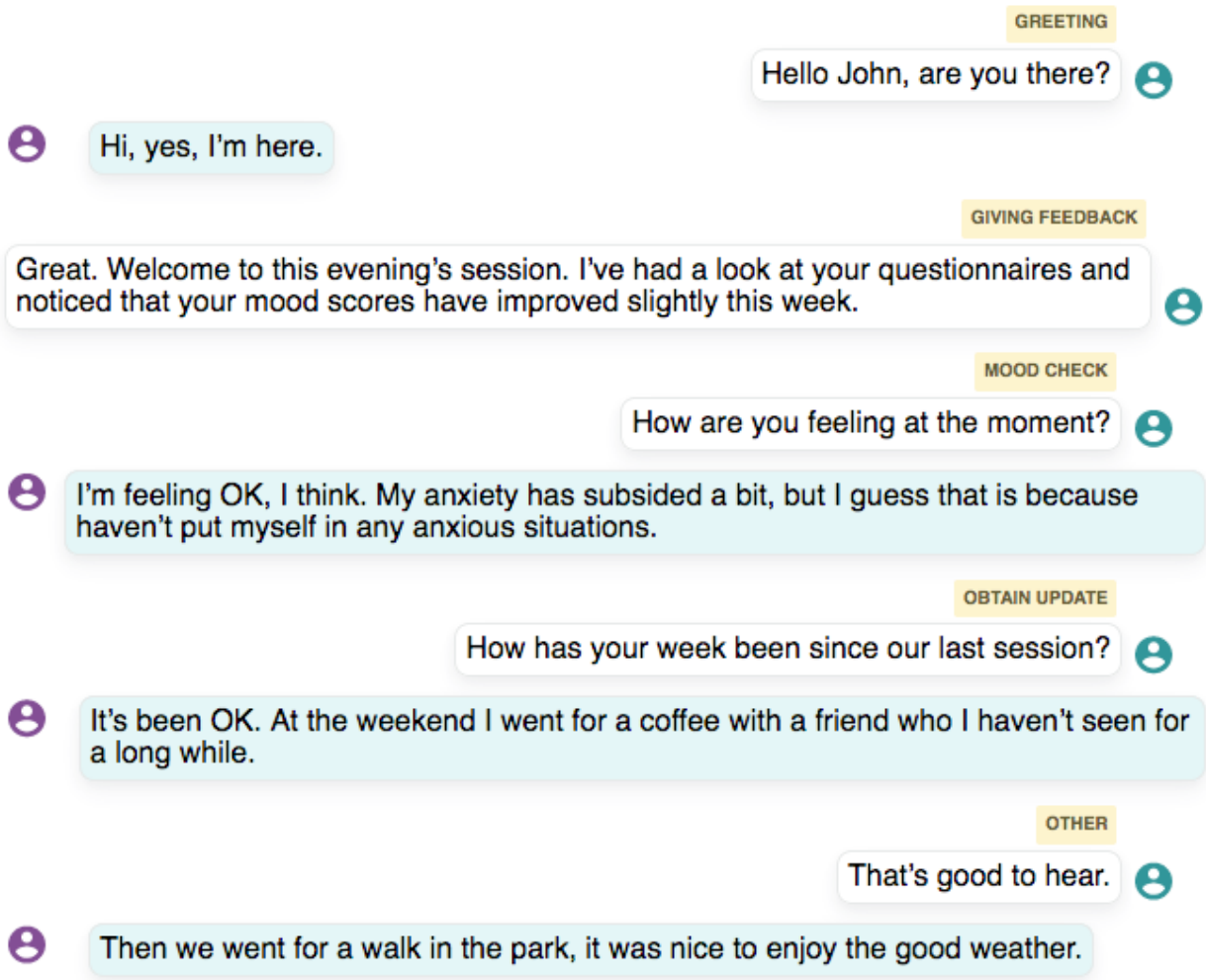
6.12.6 Developing new methods of assessing competence

The debate around the efficiency and effectiveness of tools such as the CTS-R, as discussed in section 6.2.1, above, has led to the development of new tools that assess competence. The most notable of these new tools are the Standardised Competence Rating Scale for Cognitive Therapy (Schmidt et al., 2018) and the Assessment of Core CBT Skills (Muse, McManus, Rakovishik and Kennerly, 2014). Despite these significant attempts to develop tools that more accurately reflect therapist competence and its relationship to outcome, a significant problem remains, which may explain why these two new tools have not been widely adopted. Researchers who are attempting to define competence and its' relationship with outcome are unable to obtain enough live data of therapy sessions, from a

broad range of therapists, in order to analyse and quantify the process-outcome relationship (Ewbank, Cummins, Tablan, Bateup, Catarino, Martin and Blackwell, 2019). In this respect, the IECBT method offers the opportunity to analyse, at scale, a large data set of live therapy sessions delivered to 40,000 patients[5]. The IECBT method is, in itself, a standing research trial platform, offering the opportunity to develop new tools to assess therapist competence. With the availability of data of this volume it is possible to begin to understand which therapist behaviours are closely associated with outcome. These behaviours can be quantified as items such as setting an agenda for the session, undertaking a mood check or reviewing a homework task (Ewbank et al, 2019). In addition, because of the digital nature of the treatment, it becomes possible to develop automated tools that screen, in real time, each therapy session identifying the presence, or absence of each item. This process is currently in development and is entitled 'Therapy Insights Model' (TIM). Figure 6.6 shows TIM on the Ieso Digital Health platform. Whilst the development of TIM is beyond the scope of this thesis, the concept has been directly derived from the research reported in this thesis. The findings (from this research) that higher CTS-R scores correlate with better outcomes, whilst not necessarily causal, has led to the development of an automated tool that can detect key items, or mechanisms of change that relate to outcome.[6] The implications of this on clinical practice are highly significant. This new tool (albeit still in development) could be used to provide feedback to therapists after every session, to provide data to clinical supervisors and to inform further research in this area.

[5] The number patients who had completed treatment at Ieso Digital for the period 2014-June 2019
[6] For further details of the automated tool TIM see Ewbank, Cummins, Tablan, Bateup, Catarino, Martin & Blackwell, (2019).

Figure 6.6 The Therapy Insights Model (Ewbank et al., 2019) Extract of a (fictitious) IECBT session where the utterances of a therapist are on the right, and those of the patient on the left. The utterance tag (shown in yellow) show the output of TIM



6.12.7 Therapist adherence

Outcomes-based research in the field of CBT have tended to focus on the provision of interventions that adhere to a specific protocol (Layard and Clark, 2014). The protocol follows a treatment manual and clinicians are trained to deliver treatments with adherence to the manual in research trials (Layard and Clark 2014). Once there is statistically significant evidence that the treatment protocol achieves good outcomes then the protocol is deemed efficacious and it is considered to be 'evidence-based' (Layard and Clark, 2014). Several of these evidence-based protocols are incorporated into the training of IAPT trainees (NHS England, 2018). Additionally, IAPT stipulate that High Intensity cognitive behavioural therapists are required to deliver interventions in accordance with National Institute of Health and Care Excellence (NICE) guidance using evidence-based protocols for specific disorders (NHS England, 2018). These protocols are discussed more fully in Chapter 2 of this thesis. It is impossible to know whether in fact High Intensity CBT therapists are delivering interventions that adhere to the evidence base as therapy tends to take place behind closed doors and there are very few recordings of live sessions in real-world settings (Liness, Lea, Nestler, Parker and Clark, 2016). Whilst adherence to a treatment protocol is central to the IAPT programme very little is known about whether IAPT therapists adhere to protocols after training. One study (Liness et al., 2016) suggests that therapists are less likely to adhere to a protocol and these findings support the hypotheses of Waller and Turner (2016), that CBT therapists tend to drift away from adherence to a protocol. The findings reported in this thesis would suggest that adherence is significantly related to outcomes (albeit via therapist competence) and, therefore, it would be reasonable to hypothesise that not all IAPT therapists are adhering to a protocol. This hypothesis is supported by several studies who have investigated the phenomenon known as 'therapist drift' (see Shafran, Clark, Fairburn, Arntz, Barlow, Ehlers, Freeston, Garety, Hollon, Ost, Salkovskis, Williams and Wilson, 2009, McAleavy, Castonguay and Goldfried, 2014 and Wolf and Goldfried, 2014. The term therapist drift describes an issue that is thought to occur in therapy where a therapist either omits elements of a protocol or combines several elements of different protocols, so that the

patient does not receive a treatment that is supported by an evidence base (Waller and Turner, 2016). Several reasons for therapist drift have been cited including lack of training, therapist anxiety, therapist beliefs that a particular treatment is ineffective or inappropriate for a particular patient and the therapists philosophical stance whereby some therapists may perceive CBT as an art rather than a science and this may increase therapist drift (Waller and Turner). In addition, this thesis has argued that the current training of IAPT therapists fails to adequately assess trainee competence in delivering CBT whilst adhering to a protocol, and it is hypothesised that some CBT therapists may lack knowledge in how to deliver particular protocols. One implication of the findings from this research is that therapists and their clinical supervisors may put greater emphasis on reflecting on the therapist's knowledge and ability to deliver CBT using a range of evidence-protocols. One study (Parker and Waller, 2014) has suggested that once qualified some therapists may avoid sharing their knowledge gaps with their supervisor and the supervisor may not probe the therapist's knowledge sufficiently enough to expose gaps. This latter occurrence is defined as 'supervisor drift' (Waller and Turner). Furthermore, many therapists and their supervisors have a belief that they are highly competent and that they closely adhere to the evidence base (Dennhag, Gibbons, Barber, Gallop and Crits-Christoph, 2012). The combined effect of all these issues may be further amplified where the therapist is not bringing live recordings of therapy sessions for their clinical supervisor to review. Therefore, a further implication from the findings of this study might be that IAPT services require therapists to provide recordings of therapy sessions where the patient does not appear to be improving. The emphasis here would be on the supervisor and the therapist to work together to identify whether there was therapist drift and, consequently, explore the training needs of the therapist. As with the potential implications for IAPT training programmes, these changes to clinical practice within IAPT come with an additional burden of time and resource but as each therapist receives one hour of clinical supervision each week in IAPT (NHS England, 2018) then it could be argued that this might be a good use of this time.

6.13 WIDER IMPLICATIONS

The IAPT programme currently sees approximately 900,000 patients every year and the programme has a target to increase provision of treatment to I.5 million in the year 2020-2021 (National Collaborating Centre for Mental Health, 2019). The programme was established to widen access to evidence-based psychological therapies and systemically improve outcomes (Layard and Clark 2014). Whilst the programme has been highly successful, this thesis has argued that the IAPT programme should not rest on its laurels. If the IAPT programme is going to continue to improve then it will be necessary to develop new ways of supporting therapists to deliver CBT both competently and with adherence to the evidence base.

6.13.1 Policy

NHS England's Long-Term Plan published in 2019 (NHS England, 2019) sets out a 10-year plan for builds on the NHS Five Year Forward View. The plan places a significant focus on increasing the use of digitally-enabled care throughout mental health, including improved use of centralised data and patients' records for research and development, an NHS app for use on smart phones so that patients can access information and advice more readily and the use of digital products to augment and deliver interventions in mental health.

The implications of the research reported in this thesis on health care policy are that the findings support the NHS Long-Term Plan for Mental Health to focus on patient recovery. This thesis has argued that the IECBT method affords researchers with a new way in which to conduct psychological therapy research. Given the NHS are investing in digital innovations to support service delivery, and there is a significant emphasis on learning what works for whom, then the research contained in this thesis is a timely addition to the literature in this area.

6.13.2 Embedding technology in psychological therapy

With the emphasis on embedding technology into IAPT as a way of augmenting and delivering cognitive and behavioural therapies one further implication of this research is that IAPT may choose to incorporate the methodology described in this research to monitor and assess therapist performance. This would require all IAPT therapists to deliver treatments using a method similar to IECBT. In this way, it would become possible, at even greater scale, to learn from those therapists who achieve the best outcomes and support average or below-average therapists to become better. Some studies (see Lutz, Lambert, Harmon, Tschitsaz, Schurch and Stulz, 2006, Lutz, Bohnke and Kock, 2011, Lutz, Rubel, Schiefele, Zimmerman, Bohnke and Wittman, 2015, Strauss, Lutz, Steffanowski, Wittmann, Boehnke, Rubel et al., 2015) are exploring ways in which providing feedback and/or guidance to a therapist or a patient can improve outcomes. Examples of feedback include using statistical models that reliably predict patient outcome and alert the therapist and patient when the patient is not making progress as predicted (Strauss et al.,2015, Lutz, Bohnke and Kock, 2011). Where therapy is delivered via a digital method this (either IECBT, guided self-help, virtual reality or video-conference) feedback and guidance may be delivered in a timely way, whilst the therapist is delivering a treatment session. These digitally enabled clinical decision support tools can provide feedback and guidance on good clinical practice thus enhancing, or even amplifying, the effect of evidence-based psychological therapies (Lutz, Bohnke and Kock, 2011).

6.14 CONCLUSION

This chapter has discussed the findings from this research with an emphasis on how this research has made a significant contribution to knowledge regarding the therapist variables that relate to clinical outcome in the provision of High Intensity cognitive behavioural therapy in IAPT. It has been argued that the delivery of CBT via Internet enabled CBT (IECBT) has provided a new, and currently unique, way to learn more about what drives good clinical

outcomes. It has been argued that the IECBT method represents a new research paradigm enabling researchers to study real-world data in great volume. The implications of the findings from this research on professional practice within IAPT, and also on the curriculum and assessment processes used by IAPT training programmes, has been discussed and it has been argued that the findings of this research support the conceptual framework and basic aims and objectives of IAPT. More specifically, this chapter has argued that the learning from this research would benefit therapists, clinical supervisors, IAPT training programmes, services and patients alike. Without changes in clinical practice and training, it is unlikely that the variance in outcomes between therapists and between services will change. Therefore, as IAPT continues to expand it is highly likely that not all patients will have access to interventions that give then the optimum chance of recovery.

Chapter 7 will conclude this thesis with a personal reflection of how this research has impacted on the researcher's own professional practice.

CHAPTER SEVEN: A PERSONAL REFLECTION ON THE IMPLICATIONS OF THIS RESEARCH

This thesis began with an introduction to the professional context for this research. This thesis went on to sequentially report on the literature review, methodology, findings and discussion relating to the therapist variables that are related to outcome amongst High Intensity cognitive behavioural therapists treating patients in IAPT. This final chapter concludes this thesis with a personal reflection of the process of undertaking this research and the implications of the findings from this research on the researchers own professional practice.

7.1 INTRODUCTION

I undertook this research over a period of four years whilst working, full-time, as the Chief Clinical Officer at Ieso Digital Health. It was important to me that I undertook a Professional Doctorate, rather than a PhD, because I was seeking to further develop my own skills as a Scientist Practitioner and, therefore, a direct link to my professional practice was a key focus. This commitment to career-long learning is a theme I discussed in Chapter 1 of this thesis and, as I reflect back on the last four years, I acknowledge that the process of undertaking this research had a significant impact on my practice.

7.1.1 Reflections on the process of undertaking this research

When I commenced this research in 2015, I was a CBT therapist with over thirty-years' experience of professional practice, an experienced teacher, clinical supervisor and clinical leader. Thirty-years' experience in any professional role is likely to lead to the development of skills, expertise, beliefs and assumptions about best professional practice and what works for whom (Cooke and Brown, 1999). Prior to undertaking this research I might have described this as reflexive practice or tacit knowledge, whereby I have learnt a series of

skills and expertise which are difficult for me to articulate (Cooke and Brown,1999) but, through a process of empirical enquiry, I have come to believe that I have evidence that these skills are effective. However, through this research I have found myself questioning everything about my practice and the practice of others. What I once assumed was fact and proven, I now question. For example, this research has found that fidelity to the CBT model and adherence are related to clinical outcome. My hypotheses for the research questions that relate to these findings were that there would be a relationship between these two variables and clinical outcome. Whilst the answer to both these research questions was affirmative, I was initially disappointed with the results. That part of me which is an experienced CBT therapist and a 'disciple' of evidence-based practice believed that the results from the data analysis would be statistically much stronger and that I might even go on to demonstrate a causal relationship. I realise now that this was extremely naïve. I have come to appreciate that the evidence base is only the sum of the knowledge that we have at this current time: it is the best that we have. The experience of undertaking this research has taught me that only if I am open to learning anything from my research (even that which I do not wish to find) then I may learn something that has the potential to contribute to knowledge and impact on professional practice. I hypothesise that further process-outcome research will lead to the development of new, and hitherto unknown, forms of psychological therapy that might be more effective than CBT. It is possible that the psychological therapists of the future will look back at the CBT provided by today's therapists with incredulity.

7.1.2 The 'insider researcher'

Reflection "on' and 'in' action are commonly used processes in present-day CBT (Bennett-Levy, 2006). The general aim of reflection in this context to is to enhance and hone clinical skills in order to improve patient outcomes (Bennett-Levy). In Chapter 6 of this thesis I discussed how Bennett-Levy's conceptual framework (the Declarative Procedural Reflective Model) was central to this research and in Chapter 2 I established that empirical enquiry is

fundamental to all aspects of CBT. Despite this central focus on empiricism and reflection, conducting research as an insider researcher was not without its challenges. The term 'insider researcher' is used to describe a member, or employee, of an organisation who is conducting research in their work setting (Coghlan and Brannick, 2005). The concept of a staff member conducting research in their own work setting is well-established in both education and health care settings (Costley, Elliott and Gibbs, 2013). The role of insider researcher has been described an extension of a practitioners caring role in order *"to act on behalf of another…"* (Noddings, 2003, p. 30). This aim resonates with my personal aims and objectives as Chief Clinical Officer, therapist, trainer, supervisor and researcher. Insider research in healthcare is said to be *"a formal and systematic attempt by practitioners…to understand practitioners work with the intended purpose of transforming self, colleagues and work contexts and the development of new understandings of practitioners work…"* (McCormack, 2003 p.207). It might be said that insider researchers enjoy certain advantages when conducting research in their own work setting. Firstly, insider researchers tend to have privileged access to research subjects (both practitioners and patients) expertise and the researcher is an established member of the practice team. Secondly, the insider researcher often has accumulated years of experience, knowledge and expertise that relate specifically to their practice setting (Lykkeslet-Molde and Gjengedal, 2007). I have seven years of experience in my current role but twenty-five years' experience as a cognitive behavioural therapist. This 'insider' knowledge and experience has undoubtedly led to the development of preconceptions (Lykkeslet-Molde and Gjengedal, 2007) about the delivery of CBT and my experience has influenced this research in that I had hypothesised that both therapist competence and adherence would have a strong relationship with outcome. I have come to understand that, at the start of this research, I had a strong feeling that there was a 'right' way to do CBT. This 'right' way aligned to the IAPT training curriculum. Having spent years training, supervising and managing therapists in how to deliver effective CBT, there was little doubt in my mind that fidelity to the CBT model and adherence were paramount in achieving good clinical outcomes. Whilst this research has demonstrated that there is a

statistically significant relationship between competence and adherence at the $\leq$ 0.05 level, on reflection, it is possible that my bias towards the importance of adherence and competence influenced the way I conducted this research. The process of conducting the research has forced me to question whether rating competence and adherence were the most appropriate methods for this research. Currently, I believe that it would be more fruitful to closely examine the transcripts of the most effective therapists in order to understand the mechanisms of change that these therapists employ. It might be argued that an outsider researcher, with no experience of CBT, would have been less caught up with the doctrines, professional codes and established principles of CBT and as a result they may have had a naive but more objective stance. Whilst I made attempts to remain objective, control for bias, and my own personal influence on this research, this was one of the more challenging aspects of being an insider researcher. The main issue that I contended with was my role as Chief Clinical Officer and the power and influence that this may bestow on those I work with. This may have been mediated by the fact that this research undertook a post hoc analysis of data *in silico* (see Chapter six for a description of *in silico* research). Therefore, at the time the therapy was conducted, therapists would have been unaware that their work would be studied for research purposes. However, the supervisors who rated the transcripts may have been influenced by my involvement. I have clearly articulated values about the quality of therapy provision and service delivery that I expect, and the clinical supervisors will have been exposed to these opinions. Whilst I took care to ensure that I did not provide the inter-rater reliability training or feedback to the supervisors it was impossible to completely eradicate my influence as a researcher because of the duality of roles (researcher and Chief Clinical Officer). It might be argued there are no elegant solutions to balancing the benefits of insider research with the difficulties that can be experienced in maintaining a stance of objectivity and design a study that is methodologically robust (Costley Elliott and Gibbs, 2013). Just as randomised controlled trials have been criticised for over controlling variables, so much so that the results may not generalise to real-world settings (see discussion in Chapter 3). This juxtaposition between real-world research and academic

(outsider) research was, at times, difficult as I am aware that I cannot remove myself, my influence and prior knowledge from this research. Furthermore, I became aware that there are other stakeholders in this research that I have also had to consider. Figure 7.1 shows the various stakeholders that have impacted on this research. The most controversial of which are the relationships I have with those that have power or influence over me, including my employer (Chief Executive Officer, Chair of the Board of Directors and the Board of Directors as a whole) and the founders and senior managers of the IAPT programme. I am aware that these two groups (employer and IAPT) may have competing expectations and that the results of my research may be controversial to both parties. My employer may expect my research to result in intellectual property that directly benefits the organisation and IAPT may wish me to present findings that are congruent with their agenda.

Figure 7.1 The influences and stakeholders impacting on this research (adapted from Costley, Elliott and Gibbs, 2013 p. 2)

Me

Senior clinical team and supervisors

Therapists who work at Ieso Digital Health

Wider team at Ieso Digital Health

Chief Executive Officer and Board of Directors

CBT community and key opinion leaders

Founders and senior managers of IAPT and NHS England

It might be argued that the issues I have experienced in conducting this research are merely an extension of my practice in that there are always competing demands that relate to issues such as quality versus cost effectiveness, my employer's agenda versus the IAPT agenda. I have managed these aspects of my practice, just as I managed them in my role as a researcher, in this research, in that I am acting on behalf of the patients (Noddings, 2003) in order to improve and enhance patient care. Central to this is the concept of career-long learning and through this research I have learnt from what went well and want did not.

7.1.3 What I have learned

In Chapter 1 of this thesis I outlined my commitment to career-long learning. Through the process of undertaking this research, I recognise that my primary motivation was to make a difference to others. Most notably, I wished to add to the evidence base so that other CBT therapists might be influenced by the findings of my research and, therefore, more patients would benefit. Of course, I appreciate that this thesis will, in itself, make very little difference to other CBT therapists. However, it is possible that the learning that I have derived from this research might make a difference if I disseminate it in a meaningful way. It has been suggested that the dissemination of professional learning that is perceived by the learner to lead to an enhancement in professional competence, in the caring professions, is the strongest motivator for change (Eraut, 2005). At the beginning of this research process I firmly believed that professional expertise is built on a foundation of theoretical knowledge (Spencer and Spencer, 1983). However, I now strongly suspect that theoretical knowledge alone does not equate to clinical expertise. A key example, drawn from this research, is that the IAPT training programme has based its curriculum on the acquisition of theoretical knowledge. This includes the theory that relates to each of the evidence-based treatments protocols (NHS England, 2018). Trainees are taught about each disorder specific protocol, typically in lecture style presentations. The acquisition of this theoretical knowledge aligns with the summative assessment of knowledge relating to the delivery of disorder specific

protocols (Department of Health, 2011b). CBT trainees are assessed on their ability to adhere to a protocol through their academic writing rather than a practical demonstration of the trainees' ability to adhere to a protocol when treating a patient (Department of Health 2011b). This highlights the theory-practice gap that is widely discussed in the literature relating to training health care professionals (Monaghan, 2015, Hofmann, 2013, Williams, Boyle and O'Meara, 2009, Shafran, Clark, Fairburn, Arntz, Barlow, Ehlers et al., 2009). As someone who has taught on IAPT training programmes, it had not occurred to me that (although I was aware of the concept of the theory-practice gap) trainees may not have been adequately supported to translate theoretical knowledge into practice. Whilst there is some evidence that some trainees may have a natural ability to translate theory into practice, termed "practical intelligence" (Eraut 2005 p.177), arguably some trainees will not possess this ability (Imel, Sheng, Baldwin and Atkins, 2015). This issue highlights how, before undertaking this research, I had a tendency to accept that what I had been taught (and had therefore been teaching) was the right thing to do. What I know now is that this practice is only based on the best available knowledge at that time and that I have a responsibility to question everything. Finally, I have learnt that if I am truly a Scientist Practitioner then I will reflect on the findings from this research, identify another gap in the knowledge, another research question and formulate another hypothesis in order to build on that I do not know.

7.1.4 What I have yet to learn

In section 7.1.1, above I have alluded to the fact that the most impactful of all the things that I have learned from undertaking this research is that I have so much more to learn. At the risk of sounding cliched, I have become humbled by the realisation that, whilst I have learnt a great deal in the process of undertaking this research, this learning pales in comparison with what I do not know. That is not to say that I perceive this as a negative outcome, rather more, it motivates me to continue in pursuit of new knowledge and understanding in the service of enabling therapists to be the best they can be so that more patients can get

better. This research has highlighted further areas of research that I wish to pursue, some of which I have already embarked upon. Most notably, whilst this research has established that there is a statistically significant relationship between therapist competence, adherence and clinical outcomes I am yet to establish a causal relationship. I am aware that there are likely to be other, perhaps yet unknown, therapist factors that are also related to clinical outcome. I reflect on what a privilege it is to have access to a data set that might provide the answers to these questions. Additionally, I am driven to investigate how it might be possible to improve outcomes and I question whether it will ever become possible to achieve outcomes of 100% recovery. Clearly moving towards this is a worthy goal and will involve partnership and collaboration with others using automated therapist feedback, data-driven clinical decision support tools like those being developed by Lutz, Rubel, Schiefele, Zimmermann, Bohnke and Wittmann (2015), Delgadillo, Omar and Lutz, (2016) Degadillo, Overend, Lucock, Groom, Kirby, McMillan et al., (2017).

7.2 IMPLICATIONS FOR MY OWN PROFESSIONAL PRACTICE

In Chapter 6 of this thesis I discussed the implications of the findings from this research on the professional practice of other cognitive behavioural therapists, clinical supervisors and clinical leaders. As Chief Clinical Officer of a large psychological therapy service I have overall responsibility for the quality of the therapy that is provided by a team of 619 BABCP accredited CBT therapists and the recovery rates for patients treated by these therapists. Clearly, if I have indicated that findings from this research has implications for practice for CBT therapists, clinical supervisors, and clinical services as a whole, then these same implications apply to me as a senior leader and as a therapist and clinical supervisor.

7.2.1 Implications as a senior leader and researcher

As a senior leader of a large psychological therapy service I felt a significant burden of responsibility to share the findings from this research. Whilst it was important not to infer

causality, I believe that the senior clinical team, clinical supervisors and tutors should understand the findings from this research. Once disseminated, I feel that it is my responsibility to collaborate with the team to explore the implications of the findings on the service we provide to the patients who are referred to us. This includes making a decision about how we communicate the findings to the 619 therapists working within the service and what changes we may choose to make to the clinical policies and processes employed by the service. Chapter 2 of this thesis discussed some of the policies and processes currently used as part of the Ieso method of delivering CBT. These include a significant focus on the provision of clinical supervision and continuing professional development to the therapists that work in the service. The provision of supervision and training to Ieso therapists is driven by the data collected by the service. That is to say, I place a significant emphasis on providing training and supervision that directly relates to the skill deficits of the therapists. I call this methodology 'personalised continuing professional development' in that, rather than relying solely on therapists identifying their own training needs the training and supervisory team can identify what skills individual therapists are struggling with and can direct the therapist to a relevant training module on the service's e-learning platform. Given that this research has identified that therapist competence and adherence are related to outcomes a direct impact of this finding on clinical practice is that I have supported the senior clinical team to develop specific training modules that focus on therapist competence and adherence to evidence-based protocols. I have already highlighted that it is possible that theoretical training alone may not enable all therapists to apply the theory in their clinical work with patients. If this hypothesis is true, then it will be important to explore pedagogical processes which support all therapists to bridge the theory-practice gap. In addition, it will be necessary to support clinical supervisors to identify those therapists who have lower CTS-R scores and/or are failing to use evidence-based protocols and to direct them to the appropriate training module. As part of this process, I intend to use a post hoc analysis of the outcome data for therapists who have completed 'personalised continuing professional development (CPD)' to understand whether the additional training and support made a

statistically significant difference to patient outcomes. This, in turn, leads to a further research question that relates to whether or not it is possible to teach an average or poor therapist to become a better therapist. There is currently little or no research in this area (Fairburn and Cooper, 2011).

Further research, such as exploring the efficacy of 'personalised CPD' and understanding whether it is possible to support an average therapist to become a better therapist, is another implication of the findings from the research presented in this thesis. In Chapter 6 I outlined my recommendations for future research. Whilst these recommendations were primarily aimed at an external audience they are also, in some part, my responsibility predominantly because other researchers may not have access to a large data set of therapy transcripts. Therefore, data sharing and collaborating with others, who share my curiosity in this research area, is a further personal implication of this, current, research. In the final stages of this research I purposefully sought to collaborate with researchers in this field, sharing my findings from this research and exploring future questions. Earlier in this thesis I argued that psychological research should not be focussing on top down research that places an emphasis on scientists conducting randomised controlled trials away from real-world settings. That is not to say that I do not feel that this type of research is unwarranted, I am merely arguing that there should be a greater emphasis on demonstrating the generalisability of such research in real-world settings and that there should be a greater effort made for scientists, clinicians and services to collaborate on real-world research (Strauss, Lutz, Steffanowski, Wittmann, Boehnke, Rubel et al., 2015). These reflections have resulted in research collaborations with a number of academic and clinical institutions, including the University of Exeter, Trier University, Sheffield University, Boston University and the Beck Institute. The resulting research projects are outlined in table 7.1.

Table 7.1 Ongoing research following the findings of this research

Institution	Research
University of Exeter	i) Investigating the therapist variables that relate to outcome in High Intensity therapists working with older adults ii) Investigating the efficacy of a CPD training programme for High Intensity therapists working with older adults
Trier University	i) Investigating the use of therapist feedback to improve outcome in patients who are off track, using a) static growth curve model b) a dynamic growth curve model ii) Investigating how patients and therapist may be matched based on interpersonal personality traits, in order to maximise the likelihood of recovery.
Sheffield University	i) Investigating the correlation between the formative assessment of trainee therapist competence, using the CTS-R, and an automated CTS-R using a machine learning model. ii) Investigating the processes that are related to sudden therapeutic gains. iii) Using growth curve models to support therapists and patients to address issues that relate to therapy being 'off track'.
Boston University	Investigating how adherence to a Unified Protocol relates to outcome in US therapists, delivering CBT online using the Ieso Digital Health platform
Beck Institute	Investigating the efficacy of an automated tool to rate therapist competence following face-to-face and online clinical training

In addition to the ongoing research outlined in Table 7.1, the findings from the research discussed in this thesis has been disseminated to the wider team at Ieso Digital Health and this has resulted in the development of a machine learning model that automatically assesses competence at every therapy appointment. This research is based on a model entitled the Therapy Insights Model (TIM) and has been published (see Ewbank, Cummins, Tablan, Bateup. Catarino, Martin and Blackwell, 2019). The Therapy Insights Model (see Chapter 6, figure 6.6) has been incorporated into the Ieso Digital Health Platform and the data is available to a therapist's Clinical Supervisor. The model, based on the CTS-R, indicates the presence or absence of several aspects of cognitive behavioural therapy, such as agenda setting, giving and eliciting feedback and homework setting. The model provides a score for each aspect of CBT detected for every therapy session the therapist provides. Higher scores are closely correlated with clinical outcome (Ewbank et al., 2019) and there are plans to provide therapists with their own scores in 2020. Further research exploring therapist variables and the relationship with clinical outcome is also being undertaken. This research involves the close examination of the therapy transcripts of those therapists who are consistently achieving higher outcomes in order to understand the mechanisms of change that might be related to higher outcomes.

A final implication of the research presented in this thesis as a senior clinical leader is how I use the knowledge gained to influence policy both in health care and in higher education. I have argued that the findings from this research may point to a larger problem that is endemic within IAPT training programmes and in IAPT services. This problem relates to the methods that are used to assess therapist competence and adherence. The implications of these findings, and the subsequent hypothesis that IAPT might improve recovery rates by adopting new methods of assessing therapist competence/adherence are not insignificant. The overarching aims of IAPT is to improve recovery rates, so that as many people as possible benefit from psychological therapy (Clark, Canvin, Layard, Pilling and Janecka, 2018). Therefore, as part of my role it is my responsibility to lobby senior policy makers and

those that have responsibility for the national curriculum in Higher Education for IAPT. I am currently in discussions with the board of the British Association of Cognitive and Behavioural Psychotherapy (BABCP), the Beck Institute and various universities in the United Kingdom and have presented at a number of conferences in 2017, 2018 and 2019.

7.2.2 Implications as a clinical supervisor

I continue to deliver clinical supervision to senior clinicians as part of my role as Chief Clinical Officer. The findings from the research presented in this thesis can be incorporated into my role as clinical supervisor. I can place a significant focus on enabling those that I supervise to reflect on competence and adherence in order to further test the hypothesis that this will lead to improved outcomes with their patients. This activity, in itself, has its roots in the scientist practitioner stance (Long and Hollin,1997).

7.2.3 Implications as cognitive behavioural therapist

In addition to continuing to deliver clinical supervision, I continue to practice as a CBT therapist because I believe that clinical leaders should maintain their clinical skills and should therefore continue to practice. One study suggests that leaders who fail to maintain their clinical practice become out-of-date and become less effective as a leader (see Joffe and MacKenzie-Davy, 2012). Notwithstanding this research, it is important to me to remain current in my professional practice so I can continue to learn and, therefore, enable those around me to learn. In keeping with this personal belief, the most immediate implication of this research is how I use the learning drawn from it in my own clinical practice. That would mean that I should pay greater attention to my own practical and theoretical training needs in relation to my own ability to deliver CBT with fidelity to the model, whilst adhering to an evidence-based protocol and how my own practice impacts on the clinical outcomes of the patients I treat. Additionally, I might take these self-reflections to my own clinical supervision in order that I can be guided and supported to become a better clinician.

7.3 CONCLUSION

This chapter has explored the implications of the findings from this research, and the research process as a whole, on my own professional practice. I have presented these implications in relation to how I have developed as a cognitive behavioural therapist, a supervisor, a senior clinical leader and as a researcher. I have explored the personal and professional processes that were involved in conducting this research and how these have impacted on what I have learnt and what I am currently researching.

This research has been a culmination of 5 years' work and has led to a significant contribution to the knowledge that relates to understanding which therapist variables are associated with clinical outcome in High Intensity CBT therapists working online using Internet Enabled cognitive behavioural therapy in IAPT. This research has found that therapist competence, as rated by the Cognitive Therapy Scale (CTS-R), and therapist adherence to an evidence-based treatment protocol are related to clinical outcome at the $\leq$ 0.05 level. Whilst statistically significant, this finding is not sufficient to suggest that the relationship is causal. Further research is ongoing to explore the therapy transcripts of the therapists who achieve the highest clinical outcomes in order to investigate the mechanisms of change that may be responsible for high clinical outcomes. What started out as a privileged opportunity to study therapist competence and adherence, and their relationship to clinical outcome, has turned out to be an acceptance that I (along with the research in his area) simply do not know enough about why and how CBT works. Therefore, it is necessary to answer this question in order to develop new tools that will hopefully be more effective in measuring therapist competence. This is important work because "…the patients are waiting…" (Dr. Paul Janssen, Johnson and Johnson).

REFERENCES

Aboujadoude, E., and Starcevic, V. (2014). *Mental Health in the Digital Age*, Oxford, Oxford University Press.

Almlöv, J., Carlbring, P., Berger, T., Cuijpers, P., and Andersson, G. (2009). Therapist Factors in Internet- Delivered Cognitive Behavioural Therapy for Major Depressive Disorder. *Cognitive Behaviour Therapy,* 38, 247-254.

Anderson, T., Ogles, B.M., Patterson, C.L., Lambert, M.J., Vermeersch, D.A. (2009). Therapist effects: Facilitative interpersonal skills as a predictor of therapist success. *Journal of Clinical Psychology*, 65, 755-768.

Andersson, G. 2009, "Using the Internet to provide cognitive behaviour therapy", *Behaviour research and therapy,* vol. 47, no. 3, pp. 175-180.

Andersson, G. (2014). Internet-based psychotherapy in: Aboujaoude, W. and Starcevic, V. (Eds.) *Mental Health in the Digital Age*, p. 196-219, Oxford, Oxford University Press.

Andersson, G. (2015), *The Internet and CBT: A Clinical Guide*, Boca Raton, CRC Press.

Andrews, H.B., (2000). The myth of the scientist practitioner: A reply R.King. *Australian Psychologist*, 35, 60-63.

BABCP (2010) *Standards of Conduct*. Available online at:
http://www.babcp.com/files/About/BABCP-Standards-of-Conduct-Performance-and-Ethics.pdf
Accessed 16th June 2018.

BABCP (2018) About The BABCP. Available online at:
https://www.babcp.com/About/About.aspx#1
Accessed 9th September 2019

BABCP (2019). Cognitive Behavioural Psychotherapist Full Accreditation. Available online at
https://www.babcp.com/Accreditation/CBP/CBP-Full-accreditation.aspx
Accessed 9th October 2019.

Baker, T.B., McFall, R.M. and Shoham, V. (2008). Current status and future prospects of clinical psychology: Toward a scientifically principled approach to mental and behavioural healthcare, *Psychological Science in the Public Interest,* 9 (1), 67-103.

Baker, C. (2018). *Mental Health Statistic for England: Prevalence, Services and Funding.* House of Commons Library, Number 6988, 25th April 2018.

Baldwin, S.A., and Imel, Z.E. (2013). Therapist effects: Finding and methods. In Lambert, M.J. (Ed.), *Bergin and Garfield's Handbook of Psychotherapy and Behavior Change* (6th edition, pp. 258-297). New York, Wiley.

Bandura, A., Adams, N.E., and Beyer, J. (1977). Cognitive processes mediating behavioral change. *Journal of Personality and Social Psychology,* 35, 125- 139.

Barak, A., Hen, L., Boniel-Nissim, M., & Shapira. N. (2008). A Comprehensive Review and a Meta-Analysis of the Effectiveness of Internet-Based Psychotherapeutic Interventions. *Journal of Technology in Human Services, 26 (3),* 109-160.

Barber, J.P., Liese, B.S., and Abrams, M.J. (2003). Development of the cognitive therapy adherence and competence scale. *Psychotherapy Research*, 13, 205-221.

Barkham, M., Margison, F., Leach, C., Lucock, M., Mellor-Clark, J., Evans, C., Benson, L., Connell, J., Audin, K., and McGrath, G. (2001). service profiling and outcomes benchmarking using the CORE-IM: Towards practice-based evidence in the psychological therapies. *Journal of Consulting and Clinical Psychology*, 69, 184-196.

Barkham, M., Lutz, W., Lambert, M.J., and Saxon, D. (2017). Therapist effects, effective therapists and the law of variability in Castonguay, L.G., and Hill, C.E.,(Eds.). (2017). *How and Why are Some Therapists Better Than Others? Understanding Therapist Effects*. American Psychological Association, Washington DC.

Barlow, D.H. (2006). Psychotherapy and psychological treatments: the future. *Clinical Psychology: Science and Practice*, 13 (1), 216-220.

Barlow, D.H., Hayes, S.C., and Nelson, R.M. (1984). *The Scientist Practitioner: Research and Accountability in Clinical and Educational Settings*. New York, Pergamon.

Barnfield, T.V., Mathieson, F.M., and Beaumont, G.R. (2007). Assessing the development of competence during postgraduate cognitive behavioural therapy training. *Journal of Cognitive Psychotherapy*, 21, 140-147

Batterham, P.J., Neil, A.L., Bennett, K., Griffiths, K.M., and Christensen, H. (2008). Predictors of adherence among community users of a cognitive behavior therapy website. *Patient Preference and Adherence*, 2, 97-105.

Bateup, S.E., Palmer, C., and Catarino, A. (In review). Using technology to understand how therapist variables correlate to clinical outcomes in IAPT. *Behavioural and Cognitive Psychotherapy,* In review.

Beck, A. T. (1967). *Depression: Causes and treatment*. Philadelphia: University of Pennsylvania Press.

Beck, A.T. (1970). Cognitive therapy: Nature and relation to behavior therapy. *Behavior Therapy*, 1, 184-200.

Beck, A.T. (1976). *Cognitive Therapy and the Emotional Disorders.* New York, International Universities Press.

Beck, J.S. (2011). *Cognitive Behavior Therapy: Basics and Beyond*. New York, Guildford Press.

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

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

Beck Institute for Cognitive Behavior Therapy. (2018). *Cognitive Therapy Rating Scale.* Available online at https://beckinstitute.org/get-informed/tools-and-resources/professionals/cbt-basics-and-beyond-patient-worksheets/ accessed 7th October 2019.

Bennett-Levy, J. (2005). Therapist Skills: A cognitive model of their acquisition and refinement. *Behavioural and Cognitive Psychotherapy*, 34, 57-78.

Bennett-Levy, J. (2006). Therapist Skills: A cognitive model of their acquisition and refinement. Behavioural and Cognitive Psychotherapy, 34, 57-78.

Bennett-Levy, J., Butler, G., Fennell, M., Hackman, A., Mueller, M., and Westbrook, D. (Eds) (2004). *Oxford Guide to Behavioural Experiments in Cognitive Therapy*. Oxford: Oxford University Press.

Bennett-Levy, J., and Thwaites, R. (2007). Self and self-reflection in the therapeutic relationship in Gilbert, P., and Leahy, R.L. (Eds.) *The Therapeutic Relationship in the Cognitive Behavioural Psychotherapies*. Routledge, East Sussex.

Bennett-Levy, J., and Finlay-Jones, A. (2018). The role of personal practice in therapist skill development: A model to guide therapists, educators, supervisors and researchers. *Cognitive Behaviour Therapy*, 47, 185-205.

Bennett-Levy, J. and Lee, N.K. (2012). Self-practice and self-reflection in cognitive behavioural therapy training: What factors influence trainees' engagement and experience of benefit. *Behavioural and Cognitive Psychotherapy*, 42 (11), 48-64.

Berger, T., Hammerli, K., Gubser, N., Andersson, G., and Caspar, F. (2011). Internet-based treatment of depression: A randomised controlled trial comparing guided with unguided self-help. *Cognitive Behaviour Therapy*, 40, 251-266.

Bernstein, D.A., Borkovec, T.D., and Hazlett - Stevens, H. (2000). *New directions in progressive relaxation training: A guidebook for helping professionals* . Westport, CT: Praeger Publishers

Bettany-Saltokov, J. (2012). How to do a Systematic Literature Review In Nursing: A Step-by-Step Guide, Milton Keynes, Open University Press

Beutler, L.E., Malik,, M., Alimohamed, S., Harwood, T.M., Talebi, H., Noble, S., and Wong, E. (2004). Therapist variables. In Lambert, M.J. (Ed.), *Bergin and Garfield's Handbook of Psychotherapy and Behavior Change,* (5th Edition, pp. 227-306). New York, Wiley.

Binnie, J, (2015). Do you want therapy with that? A critical account of working within IAPT. *Mental Health Review Journal*, 20, 79-83

Blackburn, I.M., Milne, D.L. James, I.A., Baker, C., Standart, S. and Garland, A. (2001) The revised cognitive therapy scale (CTS-R): Psychometric properties. *Behavioural and Cognitive Psychotherapy*, 29 (4), 431-446.

Booth, A., Sutton, A., Papaioannou, D. (2016). *Systematic Approaches to a successful Literature Review*. London, Sage.

Bordin, E.S. (1979). The generalizability of the psychoanalytic concept of the working allinace, *Psychotherapy: Theory, Research and Practice*, 16, 252-260.

Borgerson, K. 2013, "Are explanatory trials ethical? Shifting the burden of justification in clinical trial design", *Theoretical medicine and bioethics,* vol. 34, no. 4, pp. 293-308.

Bohnke, J.R., Lutz, W., and Delgadillo, J. (2014). Negative affectivity as a transdiagnostic factor in patients with common mental health disorders. *Journal of Affective Disorders*, 166, 270-278.

Borkovec, T.D., Echemendia, R.J., Rausea, S.A., and Ruiz, M. (2001). The Pennsylvania practice research network and future possibilities for clinically meaningful and scientifically rigorous psychotherapy effectiveness research. *Clinical Psychology: Science and Practice*, 8, 155-167.

Borkovec, T.D., & Sharpless, B. (2004). Generalized Anxiety Disorder: Bringing Cognitive Behavioral Therapy into the Valued Present. In S. Hayes, V. Follette, & M. Linehan (Eds.), *New directions in behavior therapy,* pp. 209 - 242. New York: Guilford Press.

Botella, C., Garcia-Palacios, A., Vizcaino, Y., Herrero, R., Banos, R.M. and Belmonte, M.A. (2013). Virtual reality in the treatment of fibromylagia: A pilot study. *Cyberpsychology, Behaviour and Social Networking*, 16 (3), 215-223.

Boulet, J.R. (2008) Summative assessment in Medicine: The promise of simulation for high stakes evaluation. Academic Emergency Medicine, 15, 1017-1024.

Branson, A., Shafran, R., and Myles, P. (2015). Investigating the relationship between competence and patient outcome with CBT. *Behaviour Research and Therapy*, 68, 19-26.

Braun, J.D., Strunk, D.R., Sasso, K.E., and Cooper, A.A. (2015). Therapist use of socratic questioning predicts session-to-session symptom change in cognitive therapy for depression. *Behaviour Research and Therapy,* 70, 32-37.

Brosan, L., Reynolds, S., and Moore, R.G. (2007). Factors associated with competence in cognitive therapists. *Behavioural and Cognitive Psychotherapy*, 35, 179-190.

Brown, L.A., Craske, M.G., Glenn, D.E., Stein, M.B., Sullivan, G., Sherbourne, C., Bystriksy, A., Welch, S., Campbell-Sills, L., Lang, A., Roy-Byrne, P., and Rose, R. (2013). CBT competence in novice therapists improves anxiety outcome. *Depression and Anxiety*, 30, 97-115.

Brown, L.A., Lambert, M., Jones, E.R., and Minami, T. (2005). Identifying highly effective therapists in a managed care environment. *The American Journal of Managed Care*, 11, 513-520.

Brown, K.W., and Ryan, R.M. (2003). The benefits of being present: Mindfulness and its role in psychological wellbeing. *Journal of Personality and Social Psychology*, 84, 822-848.

Browne, N.S., and Doshi, M. (2006). Assessing professional and clinical competence: The way forward. Advances in Psychiatric Treatment, 12, 81-91.

Bruijniks, S. J.E., Franx, G., and Huibers, M.J.H. (2018). The implementation and adherence to evidence-based protocols for psychotherapy for depression: the perspective of therapists in Dutch specialised mental health care. *BMC Psychiatry,* available online at https://bmcpsychiatry.biomedcentral.com/articles/10.1186/s12888-018-1768-x Accessed on 24th October 2019.

Burch, S., Preston, C., Bateup, S., and Hina, F. (2017). The use of internet-enabled cognitive behavioural therapy in the treatment of depression and anxiety amongst older people. *The International Journal of Aging and Society*, 8, 1-11.

Burch, S., Preston, C., and Bateup, S. (2018). True technology-enabled mental health care: Trust, agency and ageing. *mHealth*. Available online at https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6131528/ Accessed 22nd November 2019.

Campling, P. (2019). The industrialisation and marketisation of healthcare. In Jackson, C and Rizq, R. (Eds) *The Industrialisation of Care: Counselling and Psychotherapy and the Impact of IAPT*. Monmouth, PCCS Books.

Carroll, K.M, Nich, C., Sifry, R.l., Nuro, K.f., Frankforter, T.L., Ball, S.A., et al., (2000). A general system for evaluating therapist adherence and competence in psychotherapy research in the addictions. *Drug and Alcohol Dependence.* 57, 225-238.

Castonguay, L.G., Hayes, A.M., Goldfried, M.R., and DeRubeis, R.J. (1995). The focus of therapist's intervention in cognitive therapy for depression. *Cognitive Therapy and Research*, 19, 485-503.

Castonguay, L.G., and Beutler, L.E. (Eds.). (2006). *Principles of Therapeutic Change That Work.* New York, Oxford University Press.

Castonguay, L.G., Eubanks, C.F., Goldfried, M.R., Muran, J.C., Lutz, W. (2015). Research on psychotherapy integration: building on the past, looking to the future. *Psychotherapy Research*, 25, 365-382.

Castonguay, L.G., and Hill, C.E.,(Eds.). (2017). *How and Why are Some Therapists Better Than Others? Understanding Therapist Effects*. American Psychological Association, Washington DC.

Castro, V.P., Sanchez, M.J.R., Gonzalez, C.T.P, Bethencort, J.M. de la Fuente Portero, J.A. (2014) Cognitive behavioral treatment and antidepressants combined with virtual reality exposure for patients with chronic agoraphobia, *International Journal of Clinical and Health Psychology*,14 (1) 9-17.

Catarino, A., Bateup, S., Tablan, V., Inness, K., Freer, S., Richards, A., Stott, R., Hollon, S.D., Chamberlain, S.R., Hayes, A., and Blackwell, A. (2018). Demographic and Clinical predictors of response to internet-enabled cognitive-behavioural therapy for depression and anxiety. *British Journal of Psychiatry*, 4, 411-418.

Centre for Reviews and Dissemination (2008). *Systematic Reviews: CRD's Guidance for Understanding Reviews in Health Care.* Centre for Reviews and Dissemination, University of York. Available online at https://www.york.ac.uk/media/crd/Systematic_Reviews.pdf
Accessed 1st November 2019.

Chambless, D.L., and Ollendick, T.H. (2001). Empirically supported psychological interventions: Controversies and evidence. *Annual Review of Psychology*, 52, 685-716.

Chow, D.L., Miller, S.D., Seidel, J.A., Kane, R.T., Thornton, J.A., and Andrews, W.P. (2015). The role of deliberate practice in the development of highly effective psychotherapists. Psychotherapy, 52, 337-345.

Christensen, H., Griffiths, K.M., Korten, A.E., Brittliffe, K., and Groves, C. (2004). A comparison of changes in anxiety and depression symptoms of spontaneous users and trial participants of a cognitive behavior therapy website. J*ournal of Medical Internet Research.* Available online at https://www.jmir.org/2004/4/e46/
Accessed 25th October 2019.

Clark , D.M. (2005) A cognitive Perspective on Social Phobia in W. Ray Crozicr W.R. & L.L. Alden. *The Essential Handbook of Social Anxiety for Clinicians* Chichester : John Wiley & Sons .

Clark, D.M. (2011). Implementing NICE guidelines for the psychological treatment of depression and anxiety disorders: the IAPT experience. *International Review of Psychiatry*, 23, 318-327.

Clark, D. M. (2012). The English Improving Access to Psychological Therapies (IAPT) Programme: History and Progress, in McHugh, K.R. and Barlow, D.H (Eds) *Dissemination and Implementation of Evidenced-Based Psychological Interventions*, Sage, London.

Clark, D.M. (2018). Realising the mass oublic benefit of evidence-based psychological therapies: The IAPT Program. *Annual Review of Clinical Psychology*, 7, 159-183.

Clark, D.M., Canvin, L., Green, J., Layard, R., Pilling, S., and Janecka, M. (2018). Transparency about the outcomes of mental health services (IAPT approach): an analysis of public data. *The Lancet*, 391, 679-686.

Clark, D., Field, S. & Layard, R. (2012), Mental health loses out in the National Health Service, *The Lancet,* vol. 379, no. 9834, pp. 2315-2316.

Clark, D., Layard, R., Smithies, R., Richards, D.A., Suckling, R. and Wright, B. (2009) Improving access to psychological therapy: Initial evaluation of two UK demonstration sites, Behaviour Research and Therapy, 47 (4), 910-920.

Clark, D. and Wells, A. (1995) A cognitive model of social phobia in Heimberg, R., Liebowitz, M., Hope, D.A. and Schneier, F.R. (eds), *Social Phobia: diagnosis, assessment and treatment,* Guildford Press, New York.

Coghlan, D. (2007) Insider action research doctorates: Generating actionable knowledge, *Higher Education*, 54 (3), 293-306.

Coghlan, D., and Brannick, T. (2005). *Doing Action Research in Your Own Organisation*. Thousand Oaks, California, Sage Publications.

Coldwell, C.M. and Bender, W.S. (2007). The effectiveness of assertive community treatment for homeless populations with severe mental illness: A meta-analysis, *American Journal of Psychiatry*, 164, 393-394.

Collier, S. & Scally, G. 2015, "Can private providers be trusted to run NHS hospitals after Hinchingbrooke?", *BMJ (Clinical research ed.),* vol. 350, pp. 289.

Connor, K.M., Davidson, J.R., Churchil, L.E., Sherwood, A., Foa, E. and Weisler, R.H. (2000) Psychometric properties of the Social Phobia Inventory (SPIN). New self-rating scale. *British Journal of Psychiatry*, 176: 379-386

Conway, J.B. (1988) Differences among clinical psychologists: Scientists, practitioners, and scientist-practitioners. Professional Psychology: *Research and Practice*, 19, (642-655).

Cook, S.D., and Brown, J.S., (1999). Bridging epistemologies: The generative dance between organisational knowledge and organisational knowing. *Organization Science*, 10, 381-400,

Cooke, J., Ariss, S., Smith, C., and Read, J. (2015). On-going collaborative priority-setting for research activity: a method of capacity building to reduce the research-practice translational gap. *Health Research Policy and Systems.* Available online at https://health-policy-systems.biomedcentral.com/articles/10.1186/s12961-015-0014-y
Accessed 26th October 2019.

Costley, C., Elliott, G., and Gibbs, P. (2013). *Doing Work Based Research: Approaches to Enquiry for Insider-Researchers*, Sage, London.

Cowdrey, N.D., and Waller, G. (2015). Are we really delivering evidence-based treatments for eating disorders? How eating-disordered patients describe their experience of cognitive behavioral therapy. *Behaviour Research and Therapy*, 75, 72-77.

Craske, M.G., and Barlow, D.H. (2007). *Mastery of your anxiety and panic (Therapist guide)* 4th edition) Oxford: Oxford University Press

Crits-Christoph, P., Baranackie, K., Kurcias, J., Beck, A., Carroll, K., Perry, K., Luborsky, L., McLellan, A., Woody, G., Thompson, L., Gallager, D., and Zitrin, C. (1991). Meta-analysis of therapist effects in psychotherapy outcome studies. *Psychotherapy Research,* 1, 81-91.

Cuijpers, P., Donker, T., van Straten, A., and Andersson, G. (2010). Is guided self-help as effectiive as face-to-face psychotherapy for depression and anxiety disorders? A meta-analysis of comparative outcome studies. *Psychological Medicine,* 40, 1943-1957.

Cuijpers, P., Marks, I.M., van Straten, A., Cavanagh, K., Gega, L., and Andersson, G. (2009). Computer-aided psychotherapy for anxiety disorders: A meta-analytic review. *Cognitive Behaviour Therapy*. 38, 66-82.

Cuijpers, P., Weitz, E., Twisk, J., Kuehner, C., Cristea, I., David, D, DeRubeis, R.J., Dimidjian, S., Dunlop, B.W., Faramarzi, M., Heferl, U., Jarrett, R.B., Kennedy, S.H., Kheirkhah, F., Mergl, R., Miranda, J., Mohr, D.C., Segal, Z.V., Siddique, J., Simons, A.D., Vittengl, J.R., and Hollon, S.D. (2014). gender as predictor and moderator of outcome in cognitive behavior therapy and pharmacotherapy for adult depression: An individual patient data meta-analysis. *Depression and Anxiety*. 31, 941-951.

Data.gov.uk. (2016). *National Audit of Psychological Therapies*. Available online at https://data.gov.uk/dataset/c1015196-b320-4212-b4da-36e63fb9eb9b/national-audit-of-psychological-therapies-napt
Accessed 29th September 2019.

Davies, S.C. (2013). *Annual Report of the Chief Medical Officer*. London, Department of Health.

Davis, K., Grey, N., and Gould, D. (2009). What are scoping studies? A review of the nursing literature. *International Journal of Nursing Studies*, 46, 1386-1400.

Davis, M.L., Thwaites, R., Freeston, M., and Bennett-Levy, J. (2014). A measurable impact of a self-practice/self-reflection programme on the therapeutic skills of experienced cognitive-behavioural therapists. *Clinical Psychology and Psychotherapy*, 22, 176-184.

Deacon, B.J., Farrell, N.R., Kemp, J.J., Dixon, L.J., Sy, L.J., Zhang, A.R., and McGrath, P.B. (2013). Assessing therapist reservations about exposure therapy for anxiety disorders: The Therapist Beliefs about Exposure Scale. J*ournal of Anxiety Disorders*, 27, 772-780.

de Graaf, L.E., Gerhards, S.A., Arntz, A., Riper, H., Metsemakers, J.F., Evers, S.M., Severens, J.L., Widdershoven, G. and Huibers, M.J. (2009). Clinical effectiveness of online computerised cognitive behavioural therapy without support for depression in primary care: A randomised controlled trial. *British Journal of Psychiatry*, 1, 73-80.

Delgadillo, J., Gellatly, J., and Stephenson-Bellowood, S. (2015). Decision making in stepped care: How do therapists decide whether to prolong treatment or not? *Behavioural and Cognitive Psychotherapy*, 43, 328-341.

Delgadillo, J., Moreea, O., and Lutz, W. (2016). Different people respond differently to therapy: A demonstration using patient profiling and risk stratification. *Behaviour Research and Therapy*, 79, 15-22.

Delgadillo, J., Overend, K.J., Lucock, M., Groom, M., Kirby, N., Mcmillan, D., Gilbody, S., Lutz, W., Rubel. J.A., and de Jong, K. (2017). Improving the efficiency of psychological treatment using outcome feedback technology. *Behaviour Research and Therapy*, 99, 89-97.

Delgadillo, J., Saxon, D., and Barkham, M. (2018). Associations between therapists' occupational burnout and their patients' depression and anxiety treatment outcomes. Depression and Anxiety. Available online at http://eprints.whiterose.ac.uk/129198/1/Therapist_burnout_study_2018_preprint.pdf Accessed 1st November 2019.

Del Re, A.C., Fluckiger, C., Horvath, A.O., Symonds, D., and Wampold, B.E. (2012). Therapist effects in the tehrapist alliance-outcome relationship: A restricted-maximum likelihood meta-analysis. *Clinical Psychology Review*, 32, 642-649.

Demerouti, E., Bakker, A.B., Nachreiner, F., and Schaufeli, W.B. (2001). The job demnads-resources model of burnout. *Journal of Applied Psychology*, 86, 499-512.

Dennhag, I., Connolly Gibbons, M.B., Barber, J.P., Gallop, R., and Crits-Christoph, P. (2012). Do supervisors and independent judges agree on evaluations of therapist adherence and competence in the treatment of cocaine dependence? *Psychotherapy Research*, 22, 720-730.

Department of Health (2009). *Which Talking Therapy for Depression*. Available online at https://thesf.org.uk/wp-content/uploads/2018/10/Which_Talking_Therapy_For_Depression_-_print-ready_version_with_cropmark___-1.pdf
Accessed 25th October 2019.

Department of Health (2011a). *Talking therapies: a four-year plan of action. Available online* at
https://www.gov.uk/government/publications/talking-therapies-a-4-year-plan-of-action
Accessed on 9th June 2019

Department of Health (2011b). *National Curriculum for High Intensity Cognitive Behavioural Therapy Course.* Leeds, NHS Digital.

Department of Health (2015). *The Mandate: A mandate from the government to NHS England April 2015 to March 2016* available online at:
https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/486818/mndate-NHSE-15_16.pdf
Accessed 14th September 2019

DeRubeis, R.J., and Feeley, M. (1990). Determinants of change in cognitive therapy for depression. *Cognitive Therapy and Research*, 14, 469-482.

DeRubeis, R.J., Gefland, L.A., German, R.E., Fournier, J.C. and Forand, N.R. (2014). Understanding processes of change: How some patients reveal more than others- and some groups of therapists less- about what matters in psychotherapy. *Psychotherapy Research*, 24, 419-428.

Dijkers, M.P. (2009). The value of traditional reviews in the era of systematic reviewing. *American Journal of Physical Medicine and Rehabilitation*, 88, 423-430.

Dreyfus, H.L., and Dreyfus, S.E. (1986). *Mind over machine: The Power of human intuition and expertise in the era of the computer*. The Free Press, New York.

Dreyfus, H.L. (1989). *The Dreyfus model of skill acquisition in Burke, J. (Ed) Competency based education and training*. London, Falmer Press.

Dugas, M. J., and Koerner, N. (2005). The cognitive - behavioral treatment for generalized anxiety disorder: Current status and future directions. *Journal of Cognitive Psychotherapy: An International Quarterly,* 19 *,* 61 - 81

Ehlers, A., Clark, D.M., Hackmann, A., McManus, F., & Fennell, M. (2005). Cognitive therapy for PTSD: development and evaluation. *Behaviour Research and Therapy, 43* , 413 - 431.

Ehlers, A., Grey, N., Wild, Stott, R., Linness, S., Deale, A., Handley, R., Albert, I., Hackmann, A., Manley, J., McManus, F., Brady, F., Salkovskis, P., and Clark, D.M. (2013). Implementation of cognitive therapy for PTSD in routine clinical care: Effectiveness and moderators of outcome in a consecutive sample. *Behaviour Research and Therapy*, 51, 742-752.

Elhers, A., Hackmann, A., Grey, N., Wild, J., Liness, S., Albert, I., Deale, A., Stott, R., and Clark, D.M. (2014). A randomized controlled trial of 7-day intensive and standard weekly cognitive therapy for PTSD and emotion-focused supportive therapy. *American Journal of Psychiatry*, 171, 294-304.

Elkin, I. (1999). A major dilemma in psychotherapy outcome research: Disentangling therapists from therapies. Clinical Psychology Science and Practice, 6, 10-32.

Elkin, I., Shea, M.T., Watkins, J.T., Imber, S.D., Sotsky, S.M., Collins, J.F. (1989). NIMH treatment of depression collaborative research program: general effectiveness of treatments. *Archives of General Psychiatry*, 46, 971-982.

Elkin, I., Falconnier, L., Martinovich, Z., and Mahoney, C. (2006). Therapist effects in the National Institute of Mental Health Treatment of Depression Collaborative Research Program. *Psychotherapy Research*, 16, 144-160.

Elliott, J. (1991). A model of professionalism and its implications for teacher education. *British Education Research Journal,* 17 (4), 309-318.

Elliott, L. (2018) In visibility today: Gail Ward from Disabled People Against Cuts. *Medium*. Available online at https://medium.com/@lcelliott2/in-visibility-today-gail-ward-from-disabled-people-against-cuts-501be4bc7e92
Accessed 24th October 2019.

Eraut, M. (2005). Expert and expertise: Meanings and perspectives. *Learning in Health and Social Care*, 4, 173-179.

Evening Standard. (2019). *Can AI treat depression and anxiety? Using artificial intelligence in cognitive behavioural therapy for mental health issues*. Evening Standard, 6th March, 2019.

Everitt, B.S. (1977). *The Analysis of Contingency Tables*. London, Halstead Press.

Ewbank, M.P., Cummins, R., Tablan, V., Bateup, S., Catarino, A., Martin, A.J., Blackwell, A.D. (2019). Quantifying the association between psychotherapy content and clinical outcomes using deep learning. JAMA Psychiatry. Available online at https://jamanetwork.com/journals/jamapsychiatry/fullarticle/2748757
Accessed 8th November 2019.

Eysenck, H.J. (1966). Personality and experimental psychology. *Bulletin of the British Psychological Society*, 19, 1-28.

Farmer, C.C., Mitchell, K.S., Parker-Guilbert, K., and Galovski, T.E. (2017). Fidelity to the cognitive processing therapy protocol: Evaluation of critical elements. *Behavior Therapy*, 48, 195-206.

Fairburn, C.G., and Cooper, Z. (2011). Therapist competence, therapy quality and therapist training. *Behaviour Research and Therapy*, 49, 373-378.

Fairburn, C.G., Cooper, Z., Doll, H.A., O'Connor, M.E., Bohn, K., Hawker, D.M. (2009). Transdiagnostic cognitive-behavioural therapy for patients with eating disorders: A two-site trial with 60-week follow-up. *American Journal of Psychiatry,* 166, 311-319.

Firth, N., Saxon, D., Stiles, W.B. and Barkham, M. (2019). Therapist and clinic effects in psychotherapy: A three level model of outcome variability. J*ournal of Consulting and Clinical Psychology*, 87, 345-356.

Fisher, J.E. and Happell, B. (2009). Implication of evidence-based practice for mental health nursing. *International Journal of Mental Health Nursing*, 18 (3), 179-185.

Fiske, D.W. (1977). Methodological issues in research on the psychotherapist. In Gurman, A.S., and Razin, A.M. (Eds) *Effective Psychotherapy: A Handbook of Research* (pp.23-43). New York, Pergamon.

Fluckiger, C., Del Re, A.C., Wampold, B.E., Symonds, D., and Horvath, A.O. (2012). How central is the alliance in psychotherapy? A multilevel longitudinal meta-analysis. *Journal of Counselling Psychology*, 59, 10-17.

Foa, E.B., Kozak, M.J., Salkovskis, P.M., Coles, M.E. & Amir, N. (1998). The validation of a new obsessive-compulsive disorder scale: The Obsessive Compulsive Inventory (OCI). *Psychological Assessment 10, 206-214.*

Foa, E.B. & Rothbaum, B.A. (1998). *Treating the trauma of rape: Cognitive behavioral therapy for PTSD .* New York: Guilford Press

Frank, J.D., and Frank, J.A. (1991). *Persuasion and Healing: A Comparative Study of Psychotherapy.* Baltimore, John Hopkins University Press.

Freeman, D., Reeve, S., Robinson, A., Ehlers, A., Clark, D., Spanlang, B. and Slater, M. (2017). Virtual reality in the assessment, understanding, and treatment of mental health disorders. *Psychological Medicine*, 47, 2393-2400.

Freeman, D., Haselton, P., Freeman, P., Spanlang, B., Kishore, S., Albery, E., Denne, M., Brown, P., Slater, M., and Nickless, A. (2018). Automated psychological therapy using immersive virtual reality for treatment of fear of heights: a single-blind, parallel-group, randomised controlled trial. *The Lancet*. Available online at https://www.thelancet.com/journals/lanpsy/article/PIIS2215-0366(18)30226-8/fulltext Accessed 26th October 2019.

Friedi, L., and Stearn, S. (2015). Postive affect as coercive strategy: Conditionality, activation and the role of psychology in UK government workfare programmes. Medical Humanties, 41, 40-47.

Galvan, J.L., and Galvan, M.C. (2017). *Writing Literature Reviews: A Guide for Students of the Social and Behavioural Sciences.* New York and London, Routledge.

Gauraha, N. (2017). Graphical log-linear models: fundamental concepts and applications. *Journal of Modern Applied Statistical Methods*, 16, 545-577.

Gaudiano, B.A., (2013). Cognitive-behavioral therapies: Achievements and challenges. *Evidence Based Mental Health*, 11, 5-7.

Gelder, M. (2012). *New Oxford Textbook of Psychiatry*. Oxford, Oxford University Press.

Gellatly, J. (2011). *Decision making in stepped care for common mental health problems. PhD Thesis*. The University of Manchester, School of Nursing, Midwifery and Social Work. Available online at https://www.research.manchester.ac.uk/portal/files/54506935/FULL_TEXT.PDF Accessed 24th October 2019.

Gelman, A. (2004). Analysis of variance - why it is more important than ever. *Annals of Statistics*. Available online at http://www.stat.columbia.edu/~gelman/research/published/banova7.pdf Accessed 31st October 2019.

Gibbons, C.R., Stirman, S.W., De Rubeis, R.J., Newman, C.F., and Beck A.T. (2013). Research setting versus clinic setting: Which produces better outcomes in cognitive therapy for depression? *Cognitive Therapy Research*, 37, 605-612.

Gilbert, P. (2010). An introduction to compassion focussed therapy in cognitive behaviour therapy. *International Journal of Cognitive Therapy*, 3, 97-112.

Gilbody, S., Littlewood, E., Hewitt, C., Brierley, G., Tharmanathan, P., Araya, R., Barkham, M.M., Bower, P., Cooper, C.L., Gask, L., Kessler, D., Lester, H., Lovell, K., Parry, G., Richards, D.A., Andersen, P., Brabyn, S., Knowles, S., Shepherd, C.K., Tallon, D., and White, D. (2016). Computerised cognitive behaviour therapy (cCBT) as treatment for depression in primary care (REEACT trial): Large scale pragmatic randomised controlled trial. *British Medical Journal*. Available online at. https://pdfs.semanticscholar.org/8599/e90df88b88dfc80d2293afb30f398cb88c2d.pdf?_ga=2.131020488.633805187.1572196167-1358761830.1562235177 Accessed 24th October 2019.

Ginzburg, D.M., Bohn, C., Hofling, V., Weck, F., Clark, D.M.,Stangier, I. (2012). Treatment specific competence predicts outcome in cognitive therapy for social anxiety disorder. *Behaviour Research and Therapy*, 50, 747-752.

Goldberg, S.B., Rousmaniere, T., Miller, S.D., Whipple, J., Nielden, S.L., Hoyt, W.T., and Wampold, B.E. (2016). Do therapists improve with time and experience? A longitudinal analysis of outcomes in a clinical setting. *Journal of Counselling and Psychology*, 63, 1-11.

Goldsmith, L.P., Dunn, G., Bentall, R.P., Lewis, S.W., Wearden, A.J. (2015). Therapist effects and the impact of early therapeutic alliance on symptomatic outcome in chronic fatigue syndrome. *Psychological Medicine*, 45, 2365-2373.

Goodyear, R.K., Wampold, B.E., Tracey, T.J., and Lichtenberg, J.W. (2017). Psychotherapy expertise should mean superior outcomes and demonstrable improvement over time. *The Counseling Psychologist*, 45, 54-65.

Govaerts, M.J., van der Vleuten, C.P., Schuwirth, L.W., and Muijtjens, A.M. (2007). Broadening perspectives on clinical performance assessment: Rethinking the nature of in-training assessment. *Advanced Health Science Education Theory and Practice* 12, 239-260.

Grant, A., and Townend, M. (2007). Some emerging implications for clinical supervision in British mental health nursing. *Journal of Psychiatric and Mental Health Nursing*, 14, 609-614.

Grant, N., Hotopf, M., Breen, G., Cleare, A., Grey, N., Hepjul, N., King, S., Moran, P., Pariante, C.m., Wingrove, J., Young, A.H. and Tyless, A. (2014). Predicting outcome following psychological therapy in IAPT (PRMPT): a naturalistic project protocol. *BMC Psychiatry*. Available online at: https://bmcpsychiatry.biomedcentral.com/articles/10.1186/1471-244X-14-170 Accessed 15th September 2019

Griest, R., Gustafson, D.H., Stauss, F.F., Rowse, G.L., Laughren, T.P., and Chiles, J.A. (1974). Suicide risk prediction: A new approach. *Life Threatening Behavior*. 4, 212-223.

Griest, R. and Cavanagh, K. (2013). Computerised cognitive behavioural therapy for common mental health disorders, what works for whom, under what circumstances? A systematic Review, *Journal of Contemporary Psychotherapy*, 43 (4), 243-251.

Grove, W.M., Zald, D.H., Lebow, B.S., Snitz, B.E., and Nelson, C. (2000). Clinical versus mechanical prediction: A meta-analysis. *Psychological Assessment,* 12, 19-30.

Gunturi, S.B., Narayanan, R., and Khandelwal, A. (2006). In silico ADME modelling 2: Computational models to predict human serum albumin binding affinity using ant colony systems. *Bioorganic and Medicinal Chemistry*, 15, 4118-41129.

Guy, A., Thomas, R.T., Stephenson, S., and Loewenthal, D. (2011). NICE under scrutiny, the impact of the National Institute for Health and Clinical Excellence guidelines on the provision of psychotherapy in the UK. *UK Council for Psychotherapy*. Available online at http://www.easewellbeing.co.uk/downloads/nice%20report%20ukcp.pdf Accessed 25th October 2019.

Gynani, A., Shafran, R., Layard, R., and Clark, D.M. (2013). Enhancing recovery rates: Lessons from year one of IAPT. *Behaviour Research and Therapy*, 51, 597-606.

Hamilton, M. (1960). A rating scale for depression. J*ournal of Neurology, Neurosurgery and Psychiatry*, 23, 56-62.

Hans, E., and Hiller, W. (2013). Effectiveness and dropout from outpatient cognitive behavioral therapy for adult unipolar depression: A meta-analysis of nonrandomized effectiveness studies. *Journal of Consulting and Clinical Psychology*, 81, 75-88.

Hart, C. (2001). *Doing a Literature Search: A Comprehensive Guide for the Social Sciences*. London, Sage.

Hatcher, R.L., and Gillaspy, J.A. (2006). Development and validation of a revised short version of the Working Alliance Inventory. *Psychotherapy Research*, 16, 12-25.

Haug, T., Nordgreen, T., Ost, L.G., Tangen, T., Kvale, G., Hovland, O.J., Heiervang, E.R, and Havik, O.E. (2016). Working alliance and competence as predictors of outcome in cognitive behavioral therapy for social anxiety and panic disorder in adults. *Behaviour Research and Therapy,* 77, 40-51.

Hayes, S.C. (2004). Acceptance and commitment therapy, relational frame theory and the third wave of behavioural and cognitive therapies. *Behavior Therapy*, 35, 639-665.

Health and Social Care Act (2102), available online at,
http://www.legislation.gov.uk/ukpga/2012/7/contents/enacted/data.htm
Accessed 30th May 2019.

Health and Social Care Information Centre (2014) *Psychological Therapies, Annual Report on the use of IAPT services: England 2013/14 Experimental Statistics*. Available online at: http://www.hscic.gov.uk/catalogue/PUB14899/psyc-ther-ann-rep-2013-14.pdf
Accessed on 9th June 2019

Health Research Authority. (2019). *HRA Approval, NHS Health Research Authority*. Available online at https://www.hra.nhs.uk/approvals-amendments/what-approvals-do-i-need/hra-approval/
Accessed 7th November 2019.

Hedman, E., Andersson, E., Ljotsson, B., Andersson, G., Ruck, C., and Lindefors, N. (2011). Cost effectiveness of internet-based cognitive behavior therapy versus cognitive behavioral group therapy for social anxiety disorder: Results from a randomised controlled trial. *Behaviour Research and Therapy*, 49, 729-736.

Hedman, E., and Ljotsson, B. and Lindefors, N. (2012). Cognitive behaviour therapy via the Internet: A systematic review of applications, clinical efficacy and cost effectiveness. *Expert Reviews, Pharmacoecon,* 12, 745-764.

Helmy, A., Timofeev, I., Palmer, C.R., Gore, A., Menon, D.K., and Hutchinson, P.J. (2010) Hierarchical log linear analysis of blood parameters and clinical outcome following traumatic brain injury. *Acta Neurochirurgica (Wein)* 152, 953-957.

Hepgul, N., King, S., Amarasinghe, M., Breen, G., Grant, N., Grey, N., Hotopf, M., Moran, P., Pariante, C.M., Tylee, A., Wingrove, J., Young, A.H., and Cleare, A.J. (2016). Clinical characteristics of patients assessed within an Improving Access to Psychological Therapies (IAPT) service: results from a naturalistic cohort study (Predicting Outcome Following Psychological Therapy; PROMPT). *BMC Psychiatry*, available online at https://bmcpsychiatry.biomedcentral.com/articles/10.1186/s12888-016-0736-6 Accessed 24th October 2019.

Hill, C.E., and Castonguay, L.G. (2017). Therapist effects: Integration and conclusions. In Castonguay, L.G., and Hill, C.E. (Eds) *How and Why are Some Therapists Better than Others? Understanding Therapist Effects*. Washington, American Psychological Association.

Hilvert-Bruce, Z., Rossouw, P.J., Wong, N., Sunderland, M., Andrews, G. (2012). Adherence as a determinant of effectiveness of internet cognitive behavioural therapy for anxiety and depression disorders. *Behaviour Research and Therapy,* 50, 463-468.

Hofmann, S.G. (2013). Bridging the theory-practice gap by getting even bolder with the Boulder Model. *Behavior Therapy*, 44, 603-608.

Hofmann, S.G. and Smits, J.A. (2008) Cognitive-behavioral therapy for adult anxiety disorders: a meta-analysis of randomized placebo-controlled trials, *Journal of Clinical Psychiatry*, 69 (4), 621-632.

Hollis, C., Martin, J., Armani, S., Cotton, R., Denis, M. and Lewis, S. (2013). Technological innovations in mental health care, in Davies, S.C. *Annual Report of the Chief Medical Officer*. London, Department of Health.

Holmqvist, R., Philips, B., and Barkham, M. (2015). Developing practice-based evidence: Benefits, challenges and tensions. *Psychotherapy Research*, 25, 20-31.

Hool, N. (2010). BABCP Core Curriculum Reference Document. BABCP, available online at http://www.babcp.com/files/About/BABCP-Core-Curriculum-V2-190913.pdf
Accessed 24th October 2019.

Hope, D.A , Heimberg, R.G, & Turk, Cynthia, L. (2006) *Managing social anxiety: A cognitive - behavioural approach.* Oxford: Oxford University Press.

Horvath, A.O. (2006). The alliance in context: Accomplishments, challenges and future directions. *Psychotherapy: Theory, Research, Practice, Training*, 43, 258-263.

Horvath, A.O., Del Re, A.O., Fluckiger, A.C., and Symonds, D. (2011). Alliance in individual psychotherapy. *Psychotherapy,* 48, 9-16.

Horvath, A.O., and Luborsky, L. (1993). The role of the therapeutic alliance in psychotherapy, *Journal of Consulting and Clinical Psychology*, 61, 561-573.

Horsfall, J., Cleary, M., and Hunt, G.E. (2011). Developing Partnerships in Mental Health to Bridge the Research-Practitioner Gap. *Perspectives in Psychiatric Care,* 47, 6-12.

Hunt, M. (1997). *How Science Takes Stock: The Story of Meta-Analysis*. New York, Russell Sage Foundation.

Huppert, J.D., Bufka, L.F., Barlow, D.H., Gorman, J.M., Shear, M.K., and Woods, S.W. (2001). Therapists, therapist variables, and cognitive-behavioral therapy outcome in a multicenter trial for panic disorder. *Journal of Consulting and Clinical Psychology*, 69, 747-755.

Huppert, J.D., Kivity, Y., Barlow, D.H., Gorman, J.M., Shear, M.K., and Woods, S.W. (2014). Therapist effects and the outcome-alliance correlation in cognitive behavioral therapy for panic disorder with agoraphobia. *Behaviour Research and Therapy*, 52, 26-34.

IBM. (2016). Released 2016. *IBM SPSS Statistics for Macintosh*, Version 24.0. Armonk, New York, IBM Corp.

Imel, Z.E., Sheng, E., Baldwin, S.A., and Atkins, D.C. (2015). Removing very low-performing therapists: A simulation of performance-based retention in psychotherapy. *Psychotherapy*, 52, 329-336.

Institute of Medicine. (2015). *Psychosocial interventions for mental and substance use disorders: A framework for establishing evidence-based standards*. Washington, DC: National Academies Press.

Jacobsen, N.S., Martell, C.R., and Dimidjian, S. (2001). Behavioral activation treatment for depression: Returning to contextual roots. *Clinical Psychology: Science and Practice*, 8, 255-270.

JASP Team (2018). JASP (Version 0.10.0)[Computer software].

Johns, R.G., Barkham, M., Kellett, S., and Saxon, D. (2108). A systematic review of therapist effects: A critical narrative update and refinement to review. *Clinical Psychology Review,* 67, 78-93.

Joffe, M., and Mackenzie Davey, K. (2012). The problem of identity in hybrid manager: Who are medical directors? *International Journal of Leadership in Public Services*, 8, 161- 174.

Jones, B.A., Griffiths, K.M., Christensen, H., Ellwood, D., Bennett, K. and Bennett, A. (2013) Online cognitive behaviour training for the prevention of postnatal depression in at-risk mothers: A randomised controlled trial protocol, *BMC Psychiatry*, available online at http://www.biomedcentral.com/1471-244X/13/265
Accessed on 15th March 2019.

Jones, R.B. and Ashurst, E.J. (2013) Online anonymous discussion between service users and health professionals to ascertain stakeholder concerns in using e-health services in mental health, *Health Informatics Journal*, 19 (4), 281-289.

Jorm, A.F., Patten, S.B., Brugha, T.S., and Mojtabai, R. (2017). Has increased provision of treatment reduced the prevalence of common mental health disorders? Review of the evidence from four countries. *World Psychiatry*, 16, 90-99.

Kazantis, N. (2003). Therapist competence in cognitive behavioural therapies: Review of the contemporary empirical evidence. *Behaviour Change*, 20, 1-12.

Kazantis, N., Clayton, X., Cronin, T.J., Farchione, D., Limburg, K., and Dobson, K.S. (2018). The Cognitive Therapy Scale and Cognitive Therapy Scale-Revised as measures of therapist competence in cognitive behavior therapy for depression: relations with short and long-term outcome. *Cognitive Therapy and Research*, 42, 385-397.

Kazdin, A.E. (2008). Evidenced-based treatments and delivery of psychological services: Shifting our emphasis to increase impact, *Psychological Services*, 5 (3), 201-215.

Keen, A.J. and Freeston, M.H. (2008). Assessing competence in cognitive-behavioural therapy. *The British Journal of Psychiatry*, 193, 60-64.

Kessler, D., Lewis, G., Kaur, S., Wiles, N., King, M., Weich, S., Sharp, D.J., R Araya, R., Hollinghurst, S. and Peters, T. (2009) Therapist-delivered internet psychotherapy for depression in primary care: a randomised controlled trial. *Lancet*, 374 (2), 628-634.

Kinchin, I.M., Cabot, L.B. and Hay, D.B. (2008) Using concept mapping to locate the tacit dimension of clinical expertise: towards a theoretical framework to support critical reflection on teaching. *Learning in Health and Social Care*, 7 (2), 93-104.

King, R. and Deane, F.P. (2009). Evidenced based practice in mental health services: understanding the issues and supporting and sustaining implementation, in Lloyd, C., Gournay, K. and Deane F.P. *Clinical Management in Mental Health Services*, Wiley and Blackwell

King, R.J., Orr, J.A., Poulsen, B., Giacomantonio, S.G., and Haden, C. (2017). Understanding the therapist contribution to psychotherapy outcome: A meta-analytic approach. *Administration and Policy in Mental Health and Mental health Services Research*, 44, 664-680.

Kingdon, D., Tyrer, P., Seivewright, N., Ferguson, B., and Murphy, S. (1996). The Nottingham study of neurotic disorder: Influence of cognitive therapists on outcome. *British Journal of Psychiatry,* 169, 93-97.

Kingdon, D. and Wykes, T. (2013) Increased funding needed for mental health research, *British Medical Journal*, 346, 423.

Klonoff, D.C., and True, M.W. (2009). The missing element of telemedicine for diabetes: Decision support software. *Journal of Diabetes Science and Technology*, 3, 996-1001.

Kobak, K.A., Reynolds, W.M., and Griest, J.M. (1994). Computerized and clinician assessment of depression and anxiety: Respondent evaluation and satisfaction. *Journal of Personality Assessment*, 63, 173-180.

Kobak, K.A., Wolitzky-Taylor, K., Craske, M.G., and Rose, R.D. (2017). Therapist training on cognitive behavior therapy for anxiety disorders using internet-based technologies. *Cognitive Therapy Research*, 41, 252-265.

Kolb, D.A. (1984) Experiential Learning: Experience as the Source of Learning and Development. New Jersey, Prentice Hall.

Koo, T.K., and Mae, Y.L. (2016). A guideline of selecting and reporting intraclass correlation coefficients for reliability research. Journal of Chiropractic Medicine, 15, 155-163.

Kooistra, L.C., Wiersma, J.E., Ruwaard, J. Oppen, P,. Smitt, F., Lokkerbol, J., Cuijpers, P. and Riper, H. (2014) Blended vs. face-to-face cognitive behavioural treatment for major depressive in specialized mental health care: Study protocol of a randomized controlled cost effectiveness trial, *BMC Psychiatry*, available online at
http://www.biomedcentral.com/1471-244X/14/290
Accessed 15th February 2019.

Kozak, M.J. and Foa, E.B. (1997). *Mastery of Obsessive Compulsive Disorder: A cognitive behavioural approach (Therapist guide).* Oxford: Oxford University Press

Kraus, D.R., Castonguay, L., Boswell, J.F., Nordberg, S.S., and Hayes, J.A. (2011). Therapist effectiveness: Implications for accountability and patient care. *Psychotherapy Research*, 21, 267-276.

Kraus, D.R., Bentley, J.H., Boswell, J.F., Constantino, M.J., Baxter, E.E., and Castonguay, L.G. (2016). Predicting therapist effectiveness from their own practice-based evidence. *Journal of consulting and clinical psychology*, 84, 473-483.

Krause, M., Lutz, W., and Boehnke, J.R. (2011). The role of sampling in clinical trial design. *Psychotherapy Research*, 21, 243-251.

Kronke, K., Spitzer, R.L. and Williams, J.B. (2001). The PHQ-9. Journal of General *Internal Medicine*, 16, 606-613.

Kuyken, W., and Tsivrikos, D. (2009). Therapist competence, comorbidity and cognitive behavioural therapy for depression. *Psychotherapy and Psychosomatics*, 78, 42-48.

Kyratsis, Y., Ahmad, R., and Holmes, A. (2012). Technology adoption and implementation in organisations: comparative case studies of 12 English NHS Trusts. *British Medical Journal Open,* available online at https://bmjopen.bmj.com/content/2/2/e000872.long
Accessed 24th October 2019.

Kyrios, M., Nedeljkovic,, M., Moulding, R. K., B., Austin, D., Meyer, D., Ahern, C. (2014). Study protocol for a randomised controlled trial of internet-based cognitive-behavioural therapy for obsessive-compulsive disorder. *BMC Psychiatry*. Available online at https://bmcpsychiatry.biomedcentral.com/articles/10.1186/1471-244X-14-209
Accessed 25th October 2019.

Lamers, F., van Oppen, P., Comijs, H.C., Smit, J.H., Spinhoven, P., Van Balkom, A.J., Nolen, W.A., Zitman, F.G., Beekman, A.T., and Penninx, B.W. (2011). Comorbidity, patterns of anxiety and depressive disorders in a large cohort study: The Netherlands Study of Depression and Anxiety (NESDA). *Journal of Clinical Psychiatry*, 72, 341-348.

Laska, K.M., Smith, T.L., Wislocki, A.P., Minami, T., and Wampold, B. (2013). Uniformity of evidence-based treatments in practice? Therapist effects in the delivery of cognitive processing for therapy for PTSD. *Journal of Counseling Psychology*, 60, 31-41.

Law, M., Stewart, D., Letts, L., Pollock, N., Bosch, J., and Westmorland, M. (1998). *Critical Review Form, Quantitative Studies*. McMaster University.

Layard, R., Clark, D.M., Knapp, M., and Mayraz, G. (2007). Cost -benefit analysis of psychological therapy. *National Institute Economic Review*, 202, 90-98.

Layard Report (2012) *How Mental Health Loses Out in the NHS*. Available online at: http://cep.lse.ac.uk/pubs/download/special/cepsp26.pdf Accessed on 2nd January, 2018.

Layard, R. and Clark, D. (2014) *Thrive: The power of evidence-based psychological therapies*, London, Penguin Group.

Layard, R. (2017) The Economics of Mental Health, IZA World of Labour, available online at: https://wol.iza.org/articles/economics-of-mental-health/long
Accessed 4th January 2018

Lewin, K. (1951). Problems of research in social psychology. In Cartwright, D. (Ed.s), *Field theory in social science: Selected theoretical papers* (pp. 155–169). New York, NY: Harper & Row.

Liberati, A., Altman, D.G., Tetzlaff, J., Mulrow, C., Gotzsche, P.C., Loannidis, J.P., Clarke, M., Devereaux, P.J., Kleijnen, J., and Moher, D. (2009). The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: Explanation and elaboration. *Plos Medicine*, available online at https://journals.plos.org/plosmedicine/article?id=10.1371/journal.pmed.1000100 Accessed 1st November 2019.

Liness, S., Lea, S., Nestler, S., Parker, H., and Clark, D.M. (2017). What IAPT CBT therapists do after training. *Behavioural and Cognitive Psychotherapy*, 45, 16-30.

Liness, S., Beale, S., Lea, S., Byrne, S., Hirsch, C.R., and Clark, D.M. (2019). The sustained effects of CBT training on therapist competence and patient outcomes. *Cognitive Therapy and Research*, 43, 631-641.

Llewelyn, S., and Hardy, G. (2010). Process research in understanding and applying psychological therapies. *British Journal of Clinical Psychology,* 40, 1-21.

Lloyd, C. 2009. *Clinical Management in Mental Health Services.* Hoboken, Wiley.

Loewenthal, D., and Proctor, G. (2018). *Why Not CBT? Against and for CBT Revisited*. Monmouth, PCCS Books.

Lomax, J.W., Andrews, L.B., Burruss, J.W., and Moorey, S. (2005). Psychotherapy supervision. In Beck, J.S., Gabbard, G., and Holmes, J. (Eds) *Oxford Textbook of Psychotherapy*, pp. 495- 503. Oxford, Oxford University Press.

London School of Economics (2012). *How mental illness loses out in the NHS*, available online at: http://cep.lse.ac.uk/pubs/download/special/cepsp26.pdf
Accessed 30th May 2019

Long, C., and Hollin, C., (1997). The scientist-practitioner model in clinical psychology: A critique. *Clinical Psychology and Psychotherapy*, 4, 75-83.

Lowe, M.R., Bunnell, D.W., Neeren, A.M., Chernyak, Y., and Greberman, L. (2011). Evaluating the real-world effectiveness of cognitive-behavior therapy efficacy research on eating disorders: A case study from a community-based clinical setting. *International Journal of Eating Disorders*, 44, 9-18.

Lutz, W., and Barkham, M. (2015). Therapist effects. In Cautin, R., and Lilienfield, S. (Eds.), *Encyclopedia of Clinical Psychology.* New Jersey, Wiley Blackwell.

Lutz. W., Bohnke, J.R., and Kock, K. (2011). Lending an ear to feedback systems: Evaluation of recovery and non-response in psychotherapy in a German outpatient setting. *Community Mental Health Journal*, 47, 311-317.

Lutz, W., Lambert, M.J., Harmon, S.C., Tschitsaz, A., Schurch, E., and Stulz, N. (2016). The probability of treatment success, failure and duration-what can be learned from empirical data to support decision making in clinical practice? *Clinical Psychology and Psychotherapy*, 13, 223-232.

Lutz, W., Rubel, J., Schiefele, A.K., Zimmermann, D., Bohnke, J.R., and Wittmann, W.W. (2015). Feedback and therapist effects in the context of treatment outcome and treatment length. *Psychotherapy Research*, 25, 647-660.

Lutz, W. Saunders, S.M., Leon, S.C., Martinovich, Z., Kosfelder, J., Schulte. D, Grawe, K., and Tholen, S. (2006). Empirically and clinically useful decision making in psychotherapy: Differential predictions with treatment response models. *Psychological Assessment,* 18,133-141.

Lutz, W., Rubel, J., Schiefele, A.K., Zimmermann, D., Bohnke, J.R., and Wittmann, W.W. (2015). Feedback and therapist effects in the context of treatment outcome and treatment length. *Psychotherapy Research*, 25, 647-660.

Maimaris, C., Summers, C.L., Browning, C. and Palmer, C.R**.** (1994) Injury patterns in cyclists attending an accident and emergency department: a comparison of helmet wearers and non-wearers. *British Medical Journal*, 308, 1537-1540.

Mann, C.C. (1994). Can meta-analysis make policy? *Science*, 266, 960-962.

Lykkeslet-Molde, E., Gjengedal, E., (2007). Methodological problems associated with practice-close research. *Qualitative Health Research*, 15, 699-704.

Manning, T., Sleator, R.D., and Walsh, P. (2012). Naturally selecting solutions: The use of genetic algorithms in bioinformatics. *Bioengineered*, 4, 266-278.

Mansell, W. (2008). The seven C's pf CBT: A consideration of the future challenges for cognitive behaviour therapy. *Behavioural and Cognitive Psychotherapy*,36, 641-649.

Marks, I.M. (2013). *Professor Isaac Marks talks about online CBT.* Interview Series CCBT Limited. Available online at https://www.youtube.com/watch?v=GvCHFmwwwAY Accessed 25th October 2019

Marks, I.M., Shaw, S., and Parkin, R. (1998). Computer-aided treatments of mental health problems. *Clinical Psychology: Science and Practice*, 5, 151-170.

Marks, I.M., Kenwright, M., McDonough, M. Whitaker, M., and Mataix-Cols, D., (2004). Saving clinicians time by delegating routine aspects of therapy to a computer: A randomised controlled trial in phobia/panic disorder. *Psychological Medicine*, 34, 9-18.

Martel, C.R., Addis, M.E., and Jacobson N.S. (2001). *Depression in Context; Strategies for Guided Action.* New York: WW Norton

Martindale, C. (1978). The therapist-as-fixed-effect fallacy in psychotherapy research. *Journal of Consulting and Clinical Psychology, 46*, 1526-1530.

Martinez-Mesa, J., Gonzalez-Chica, D.A., Bastos, J.L. Bonamigo, R.R., and Duquia, R.P. (2014). Sample size: How many research participants do I need in my research? *Anais Brasileiiros De Dermatologia*, 89, 609- 615.

Maas, C.J., and Hox, J.J. (2004). Sufficient sample sizes for multilevel modeling. *Methodology,* 1, 86-92.

Mathieson, F.M., Barnfield, T., and Beaumont, G. (2009). Are we as good as we think we are? Self-assessment versus other forms of assessment of competence in psychotherapy. *The Cognitive Behavioural Therapist*, 2, 43-50.

McAleavey, A.A., Castonguay, L.G., and Goldfried, M.R. (2014). Clinical experiences in conducting cognitive-behavioral therapy for social phobia. *Behavior Therapy*, 45, 21-35.

McCormack, B. (2003). Knowing and auctioning- a strategic practitioner focussed approach to nursing research and practice development. *Nursing Times Research* 18, 301-310.

McGraw, K.O., and Wong, S.P. (1996). Forming inferences about some intraclass correlation coefficients. *Psychological Methods*, 1, 30-46.

McHugh, K.R. and Barlow, D.H. (2012). Dissemination and Implementation of Evidenced-Based Psychological Interventions. *Oxford University Press*, Oxford.

McHugh, M.L. (2012). Interrater reliability: The kappa statistic. *Biochemica Medica*, 22, 276-282.

McManus, F., Rakovshik, S., Kennerley, H., Fennell, M., and Westbrook, D. (2012). An investigation of the accuracy of therapists' self-assessment of cognitive behavioural therapy skills. *British Journal of Clinical Psychology*, 51, 292-306.

McManus, F., Westbrook, D., Vazquez-Montes, M., Fennel, M., and Kennerley, H. (2010). An evaluation of the effectiveness of diploma-level training in cognitive behaviour therapy. *Behaviour Research and Therapy*, 48, 1123-1132.

McPherson, S., Evans, C.H., and Richardson, P.H. (2009). The NICE guidelines and the recovery model: Is there an evidence base for IAPT? *Journal of Mental Health,* 18, 405-414.

Miller, G.E. (1990). The assessment of clinical skills/competence/performance. *Academic Medicine*, 65, 63-67.

Miller, S.D., Duncan, B.L., Sorrell, R., Brown, G.S., and Chalk, M.B. (2006). Using outcome to inform therapy practice. *Journal of Brief Therapy*, 5, 5-22.

Milne, D.L. (2008). CBT supervision: From reflexivity to specialism. *Behavioural and Cognitive Psychotherapy*, 36, 779-786.

Milne, D., and Paxton, R. (1998). A psychological re-analysis of the scientist-practitioner model. *Clinical Psychology and Psychotherapy*, 5, 216-230.

Milne, D.L., and James, I. (2000). A systematic review of effective cognitive behavioural supervision. *British Journal of Clinical Psychology*, 46, 437-447.

Minonne, G.A. (2008). Therapist Adherence, Patient Alliance and Depression Change in the NIMH Treatment for Depression Collaborative Research Program (Doctoral dissertation. Available from ProQuest Dissertations and Theses database.

Mollon, P., (2009). The NICE guidelines are misleading, unscientific, and potentially impede good psychological care and help. *Psychodynamic Practice,* 15, 9-24.

Mollon, P. (2010). Our rich heritage-are we building upon it or destroying it? Some malign influences of clinical psychology upon psychotherapy in the UK. *Psychodynamic Practice,* 16, 7-24.

Monaghan, T. (2015). A critical analysis of the literature and theoretical perspectives on theory-practice gap amongst newly qualified nurses with the United Kingdom. *Nurse Education Today*, 35, 1-7.

Morina, N., Ijntema, H., Meyerbroker, K., and Emmelkamp, P. (2015). Can virtual reality exposure therapy gains be generalized to real-life? A meta-analysis of studies applying behavioural assessments. *Behaviour Research and Therapy*, 74, 18-24.

Moritz, S., Schilling, L., Hauschildt, M., Schroder, J., and Treszl, A. (2012). A randomised controlled trial of internet-based therapy in depression. *Behaviour Research and Therapy*, 50, 513-521.

Mullins, K. (2015). Opportunities to work in a range of IAPT accredited therapies. Available online at *https://www.healthcareconferencesuk.co.uk/.../13.2.2015_Opportunities_to_work_in_... Accessed on 25th January 2019*

Mundt, J.C., Marks, I.M., Shear, M.K., and Greist, J.H. (2012). The work and social adjustment scale: A simple measure of impairment in functioning. *British Journal of Psychiatry*, 180, 461-464.

Muse, K., and McManus, F. (2013). A systematic Review of Methods for Assessing Competence in Cognitive Behavioural Therapy. Behavioural *and Cognitive Psychotherapy, 41, (1) 4-44.*

Muse, K., McManus, F., Rakovishik, S., and Kennerley, H. (2014). *Assessment of Core CBT Skills (ACCS): User Manual*. Available online at https://www.accs-scale.co.uk/ACCS%20User%20Manual%20V3%20Final.pdf
Accessed 7th November 2019.

Muse, K., McManus, F., Rakovishik, S., and Thwaites, R. (2017). Development and psychometric evaluation of the Assessment of Core CBT Skills (ACCS): An observation-based tool for assessing cognitive behavioural competence. *Psychological Assessment*, 29, 542-555.

Nathan, P.E. and Gorman, J.M. (2007) *A Guide to Treatments that Work*, Oxford University Press, New York.

National Collaborating Centre for Mental Health. (2019). *Advancing Mental Health Equality*. National Collaborating Centre for Mental Health. Available online at https://www.rcpsych.ac.uk/docs/default-source/improving-care/nccmh/amhe/amhe-resource.pdf
Accessed 9th November 2019.

Nardini, C. (2014) The ethics of clinical trials, *ecancermedicalscience*, available online at http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3894239/
Accessed on 15th May 2015.

Naylor, C., Curry, N., Holder, H., Ross, S., Marshall, L. and Tait, E. (2013) *Clinical Commissioning Groups: Supporting Improvement in General Practice*. The Kingsfund. Available online at
http://www.kingsfund.org.uk/sites/files/kf/field/field_publication_file/clinical-commissioning-groups-report-ings-fund-nuffield-jul13.pdf
Accessed on 9th June 2015

Newnham, E.A. & Page, A.C. 2010, "Bridging the gap between best evidence and best practice in mental health", *Clinical psychology review,* vol. 30, no. 1, pp. 127-142.

NHS Confederation (2011). *Any Qualified Provider: A Discussion Paper*,

Available online at:

http://www.nhsconfed.org/~/media/Confederation/Files/Publications/Documents/Any_qualified_provider.pdf

NHS England (2015) Adult IAPT Workforce Census Report available online at

https://www.england.nhs.uk/mentalhealth/wp-content/uploads/sites/29/2016/09/adult-iapt-workforce-census-report-15.pdf

Accessed 5th January 2019

NHS Digital (2016). Psychological Therapies: Annual Report on the use of IAPT services-England, 2015-2016. Available online at https://digital.nhs.uk/data-and-information/publications/statistical/psychological-therapies-annual-reports-on-the-use-of-iapt-services/annual-report-2015-16

Accessed 26th October 2019.

NHS Digital (2018 a). Psychological Therapies: Annual report on the use of IAPT services England, further analysis on 2016-2017. Available online at:

https://digital.nhs.uk/data-and-information/publications/statistical/psychological-therapies-annual-reports-on-the-use-of-iapt-services/annual-report-2016-17-further-analyses

Accessed 16th September 2019

NHS Digital (2018b). Improving Access to Psychological Therapies (IAPT) Executive Summary (October 2017). Published January 2018. Available online at:

https://digital.nhs.uk/data-and-information/publications/statistical/psychological-therapies-report-on-the-use-of-iapt-services/psychological-therapies-reports-on-the-use-of-iapt-services-england-october-2017-final-including-reports-on-the-integrated-services-pilot-and-quarter-2-2017-18-data

Accessed 16th September 2019

NHS Digital (2019a) *Improving Access to Psychological Therapies Data Set*. NHS Digital Available online at >https://digital.nhs.uk/data-and-information/data-collections-and-data-sets/data-sets/improving-access-to-psychological-therapies-data-set
Accessed 22nd November 2019

NHS Digital (2019b) *Improving Access to Psychological Therapies (IAPT) Summary*. NHS Digital. Available online at https://digital.nhs.uk/about-nhs-digital/our-work/keeping-patient-data-safe/gdpr/gdpr-register/improving-access-to-psychological-therapies-iapt
Accessed 22nd November 2019

NHS Digital (2019c). *Directions and Data Provision Notices. NHS Digital*. Available online at https://digital.nhs.uk/about-nhs-digital/corporate-information-and-documents/directions-and-data-provision-notices
Accessed 22nd November 2019

NHS Digital (2019d). *DCB3058 Compliance with National Data Opt-outs*. NHS Digital. Available online at https://digital.nhs.uk/data-and-information/information-standards/information-standards-and-data-collections-including-extractions/publications-and-notifications/standards-and-collections/dcb3058-compliance-with-national-data-opt-outs
Accessed 22nd November 2019

NHS Digital (2019e). DCB1520 Improving Access to Psychological Therapies Data Set. NHS Digital. Available online at https://digital.nhs.uk/data-and-information/information-standards/information-standards-and-data-collections-including-extractions/publications-and-notifications/standards-and-collections/dcb-1520-improving-access-to-psychological-therapies-data-set
Accessed 6th December 2019

NHS Digital (2019f). Psychological therapies, annual report on the use of IAPT services 2018-19. NHS Digital. Available online at https://digital.nhs.uk/data-and-information/publications/statistical/psychological-therapies-annual-reports-on-the-use-of-iapt-services/annual-report-2018-19
Accessed 9th December 2019.

NHS England (2016). 2015 Adult IAPT Workforce Census Report. Available online at: https://www.england.nhs.uk/mentalhealth/wp-content/uploads/sites/29/2016/09/adult-iapt-workforce-census-report-15.pdf
Accessed 9th September 2019

NHS England (2018). Adult Improving Access to Psychological Therapies Programme, available online at:
https://www.england.nhs.uk/mental-health/adults/iapt/
Accessed 9th September 2019

NHS England (2019). NHS Long Term Plan. NHS England. Available online at https://www.longtermplan.nhs.uk/wp-content/uploads/2019/08/nhs-long-term-plan-version-1.2.pdf
Accessed 7th November 2019.

NHS Long Term Plan (2019). *The NHS Long Term Plan*. Available online at www.longtermplan.nhs.uk
Accessed 25th October 2019.

NHS Mandate (2014) *A Mandate from the Government to NHS England: April 2014 to March 2015*. Available online at: https://www.gov.uk/government/publications/nhs-mandate-2014-to-2015 Accessed on 4th April 2015.

NICE (2004a) *Anxiety: Management of anxiety (panic disorder, with or without agoraphobia, and generalised anxiety disorder) in adults in primary , secondary and community care* (Clinical Guidance 22) London, UK.

NICE (2004b) *Depression: Management of depression in primary and secondary care* (Clinical Guide 23), London, UK.

NICE (2005a) *Obsessive compulsive disorder (OCD): Core interventions in the treatment of obsessive compulsive disorder and body dysmorphic disorder,* (Clinical Guideline 31), London, UK.

NICE (2005b) *Post traumatic stress disorder (PTSD): The management of PTSD in adults and children in primary and secondary care*, (Clinical Guidline 26), London, UK.

NICE (2006) Computerised cognitive behaviour therapy for depression and anxiety (Technology Appraisal 97), London, UK.

NICE (2009a) Depression: Treatment and management of depression in adults (Clinical Guideline 90), London, UK.

NICE (2009b) Depression in adults with a chronic physical health condition: Treatment and management (Clinical Guideline 91), London, UK.

NICE (2011 a) Common mental health disorders: Identification and pathways to care, (Clinical Guideline 123), London, UK.

NICE (2011b). Commisioning Guide for people with common mental health disorders. National Institute for health and Clinical Excellence. Available online at: https://www.uea.ac.uk/documents/246046/11991919/commissioning-stepped-care-for-people-with-mental-health-disorders.pdf/5245bd72-8769-438f-8603-50c2f635b484 Accessed 15th September 2019

NICE (2011c). Commissioning stepped care for people with common mental health disorders. National Institute for Health and Social Care Excellence. Available online at https://www.uea.ac.uk/documents/246046/11991919/commissioning-stepped-care-for-people-with-mental-health-disorders.pdf/5245bd72-8769-438f-8603-50c2f635b484 Accessed 24th October 2019

NICE (2018). Developing NICE guidelines, the manual. *National Institute of Health and Social Care Excellence*. Available online at https://www.nice.org.uk/Media/Default/About/what-we-do/NICE-guidance/NICE-guidelines/Updated-Developing-NICE-guidelines-manual-for-consultation.pdf Accessed 24th October 2019.

Nissen-Lie, H.A., Goldberg, S.B., Hoyt, W.T., Falkenstrom, F., Holmqvist, R., Nielsen, S.L., and Wampold, B.E. (2016). Are therapist uniformly effective across patient outcome domains? A study on therapist effectiveness across two different treatment contexts. *Journal of Counseling Psychology*, 63, 367-378.

NMC (2015) *Professional Standards of Practice and Behaviour for Nurses and Midwives*. Available online at:
http://www.nmc.org.uk/globalassets/sitedocuments/nmc-publications/revised-new-nmc-code.pdf
Accessed 16th June 2015

Noddings, N. (2003). *Caring: A Feminine Approach to Ethics and Moral Education*. Berkeley, California, University of California Press.

No Health Without Mental Health (2011). available online at:
https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/213761/dh_124058.pdf
Accessed on 28th April 2019

OECD (2012). Social expenditure data base. *Organisation for Economic Co-operation and Development.* Available online at https://www.oecd.org/social/expenditure.htm
Accessed 25th October 2019
OECD (2018). Public spending on education (indicator). *Organisation for Economic Co-operation and Development.* Available online at:
https://data.oecd.org/eduresource/public-spending-on-education.htm (Accessed on 04 January 2019)

Okiishi, J., Lambert, M.J, Nielsen, S.L., and Ogles, B.M. (2003) Waiting for supershrink: An empirical analysis of therapist effects. Clinical Psychology and Psychotherapy, 10, 361-373.

Padesky,C.A. (1996). Developing cognitive therapist competency: Training and supervision models. In Salkovskis, P.M.(Ed.) *Frontiers of Cognitive Therapy*, pp. 266-292. New York, Guildford Press.

Papaioannou, D., Sutton, A., Carroll, C., Booth, A., and Wong, R. (2010). Literature seraching for social science systematic reviews: Consideration of a range of search techniques. *Health Information and Libraries Journal,* 2, 114-122.

Parker, Z.J., and Waller, G., (2015). Factors related to psychotherapists self-assessment when treating anxiety and other disorders. *Behaviour Research and Therapy*, 66, 1-7.

Persons, J.B., Bostrum, A., Bertagnolli, A. (1999). Results of randomized controlled trials of cognitive therapy for depression generalise to private practice. *Cognitive Therapy and Research,* 23, 535-548.

Persons, J..B., Roberts, N.A., Zalecki, C.A. and Brechwald, W.A.G. (2006), Naturalistic outcome of case-formulation driven cognitive-behaviour therapy for anxious depressed outpatients. *Behaviour Research and Therapy*, 44, 1041- 1051.

Pereira, J.A., Barkham, M., Kellett, S., and Saxon, D. (2017). The role of practitioner resilience and mindfulness in effective practice: A practitioner-based feasibility study. *Administration and Policy in Mental Health and Mental Health Services Research,* 44, 691-704.

Pilecki, B., and McKay, D., (2013). The theory-practice gap in cognitive-behavior therapy. *Behavior Therapy*, 44, 541-547.

Premier, M. and Talbot, F. (2013) CBT guided self help compares favourably to gold standard therapist-administered CBT and shows unique benefits over traditional treatment, *Behaviour Change*, 30 (4) 227-240.

Pybis, J., Saxon, D., Hill, A., and Barkham, M. (2017). the comparative effectiveness and efficiency of cognitive behaviour therapy and generic counselling in the treatment of depression: Evidence from the 2nd UK national audit of psychological therapies. *BMC Psychiatry*. Available online at https://bmcpsychiatry.biomedcentral.com/articles/10.1186/s12888-017-1370-7 Accessed 26th October 2019.

Quarmby,L., Rimes, K., Deale, A., Wessley, S. and Chalder, T. (2007). Cognitive behvaiour therapy for chronic fatigue syndrome: comparison of outcomes within and outside the confines of a randomised control trial. *Behaviour Research and Therapy*. 45, 1085 - 1094.

Raimy, V.C. (1950). *Training in Clinical Psychology*. Englewood Cliffs, New Jersey, Prentice Hall.

Raudenbush, S.W., and Bryk, A.S. (2002). *Hierarchical Linear Models: Applications and Data Analysis Methods*. Thousand Oaks, California, Sage Publications.

Rector, N.A., and Cassin, S.E. (2010). Clinical expertise in cognitive behvaioural therapy: Definition and pathways to acquisition. *Journal of Contemporary Psychotherapy*, 40, 153-161.

Reichelt, F.K., James, I.A., and Blackburn, I.M. (2003). Impact of training on rating competence in cognitive therapy. *Behavior Therapy and Experimental Psychiatry,* 34, 87-99.

Resko, S.M., Walton, M.A., Chermack, S.T., Blow, F.C., and Cunningham, R.M. (2012). Therapist competence and treatment adherence for a brief intervention addressing alcohol and violence among adolescents. *Journal of Substance Abuse Treatment,* 24, 429-437.

Resick, P.A., Monson, C.M. and Chard, K.M. (2007) *Cognitive processing therapy: Veteran/Military version* Washington, DC: Department of Veterans' Affairs

Richards, D.A., and Suckling, R. (2009). Improving access to psychological therapies: Phase IV prospective cohort study. *British Journal of Clinical Psychology*, 48, 377-396.

Richards, D., Tumulak, L. and Hevey, D. (2013). A comparison of two online cognitive behavioural interventions for symptoms depression in a student population: The role of therapist responsiveness, *Counselling and Psychotherapy Research*, 13 (3), 184-193.

Ricks, D.F. (1974). Supershrink: methods of a therapist judged successful on the basis of adult outcomes of adolescent patients. In Ricks, D.F. Roff, M. and Thomas, A. (Eds.) *Life History Research in Psychopathology*, Minneapolis, University of Minnesota Press.

Rogers, C.R. (1975). Empathic: An unappreciated way of being. *The Counseling Psychologist,* 5, 2-10.

Roth A., and Fonagy, P. (2005) *What Works for Whom*: A critical review of psychotherapy research. Second Edition. The Guildford Press, London.

Roth, A.D, and Pilling, S. (2007). The competencies required to deliver effective cognitive and behavioural therapy for people with depression and with anxiety disorders. Department of Health, available online at
https://www.researchgate.net/publication/237549113_The_Competences_Required_to_Deliver_Effective_Cognitive_and_Behavioural_Therapy_for_People_with_Depression_and_with_Anxiety_Disorders
Accessed 7th October 2019.

Roth, A.D., and Pilling, S. (2008). Using an evidenced based methodology to identify the competencies required to deliver effective cognitive and behavioural therapy for depression and anxiety disorders. *Behavioural and Cognitive Psychotherapy*, 36 (3), 129-147.

Roth, A.D., Pilling, S., and Turner, J. (2010). Therapist training and supervision in clinical trials: Implications for clinical practice. B*ehavioural and Cognitive Psychotherapy,* 38, 291-302.

Sarantakos, S. (2005). *Social Research*. New York, Palgrave Mac-Millan.

Saxon, D., and Barkham, M. (2012). Patterns of therapist variability: Therapist effects and the contribution of patient severity and risk. *Journal of Consulting and Clinical Psychology*, 80, 535-546.

Saxon, D., Barkham, M., Foster, A., and Parry, G. (2016). The contribution of therapist effects to patient dropout and deterioration in the psychological therapies. *Clinical Psychology and Psychotherapy*, 24, 575-588.

Saxon, D., Firth, N., Barkham, M. (2017). The relationship between therapist effects and therapy delivery factors: Therapy modality, dosage, and non-completion. *Administration and Policy in Mental Health and Mental Health Services Research*, 44, 705-715.

Schiefele, A.K., Lutz, W., Barkham, M., Rubel, J., Saxon, D., Schulte, D., Lambert, M.J. (2016). The extent of therapist effects in routine care- An investigation of contributing factors on 50,000 cases. *Administration and Policy in Mental Health and Mental Health Services Research*, 44, 598-613.

Schmidt, I.D., Strunk, D.R., DeRubeis, R.J., Conklin, L.R. and Braun, J.D. (2018). Revisiting how we assess therapist competence in cognitive therapy. *Cognitive Therapy and Research*, 42, 369-384.

Schon, D.A. (1983). *The Reflective Practitioner,* Temple Smith, London.

Schotke, H., Fluckiger, C., Goldberg, S.B., Eversmann, J., Lange, J. (2016). Predicting psychotherapy outcome based on therapist interpersonal skills: A five-year longitudinal study of a therapist assessment protocol. *Psychotherapy Research,* 27, 642-652.

Schulz, K.F., Altman, D.G., and Moher, D. (2010). CONSORT 2010 statement: Updated guidelines for reporting parallel group randomised trials. *BMC Medicine.* Available online at https://bmcmedicine.biomedcentral.com/articles/10.1186/1741-7015-8-18?report=reader Accessed 7th November 2019.

Scott, M.J. (2018). Improving Access to Psychological Therapies (IAPT)- the need fo radical reform. *Journal of Health Psychology*, 23, 1136-1147.

Sedgwick, P. (2012). The hawthorne effect. *British Medical Journal.* Available online at https://www.bmj.com/content/344/bmj.d8262
Accessed 7th November 2019.

Seekles, W., van Straten, A., Beekman, A., van Marwijk, H. and Cuijpers, P. (2011). Stepped care treatment for depressiona nd anxiety in primary care: a randomised controlled trial. *Trials*, available online at https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3152524/ Accessed 24th October 2019.

Sekaran, U., and Bougie, R. (2016). *Research Methods for Business: A Skill Building Approach*. Chichester, West Sussex, John Wiley and Sons

Shadish, W.R., Cook, T.D., and Campbell, D.T. (2002). *Experimental and Quasi-experimental Designs for Generalised Causal Inference*. Boston, Houghton-Mifflin.

Shafran, R., Clark, D.M., Fairburn, C.G., Arntz, A., Barlow, D.H., Ehlers, A., Freeston, M., Garety, P.A., Hollon, S.D., Ost, L.G., Salkovskis, P.M., Williams, J.M.G., and Wilson, G.T. (2009). Mind the gap: Improving the dissemination of CBT. *Behaviour Research and Therapy*, 47, 902-909.

Shaw, B.F.M., Elkin, I., Yamaguchi, J.,Olmsted, M., Vallis, T.M., Dobson, K., Lowery, A., Sotsku, S.M., Watkins, J.T., and Imber, S.D. (1999). Therapist competence ratings in relation to clinical outcome in cognitive therapy of depression. *Journal of Consulting and Clinical Psychology*, 67, 837-846.

Shealy, A.E. (1972). Combining behavior therapy and cognitive therapy in treating homosexuality. *Psychotherapy: Theory Research and Practice,* 9, 221-222.

Shieh, G. (2017). Power and sample size calculations for contrast analysis in ANCOVA. *Multivariate Behavioral Research,* 1, 1-11.

Silverman, W.K., Pina, A.A. and Viswesvaran, C. (2008). Evidenced based psychosocial treatments for phobic and anxiety disorders in children and adolescents, J*ournal of Clinical Child and Adolescent Psychology*, 37 (2), 105-130.

Simmons, A.M., Milnes, S.M., and Anderson, D.A. (2008). Factors influencing the utilization of empirically supported treatments for eating disorders. *Eating Disorders,* 16, 342-354.

Simmons, M., (2007). Insider ethnographer: Tinker, tailor, researcher or spy? *Nurse Researcher*, 14, 7-17.

Smith, M.L., and Glass, G.V. (1977). Meta-analysis of psychotherapy outcome studies. *American Psychologist*, 32, 752-760.

Smith, M.L., Glass, G.V., and Miller, T.I. (1980). T*he Benefits of Psychotherapy*. Baltimore, The John Hopkins University Press.

Smits, N., Paap, M.C., and Bohnke, J.R. (2018). Some recommendations for developing multidimensional computerized adaptive tests for patient-reported outcomes. *Quality of Life Research*, 27, 1055-1063.

So, M., Yamaguchi, S., Hashimoto,S., Sado, M., Furukawa, T.A. and McCrone, P. (2013). Is computerised CBT really helpful for adult depression? A meta analytic re-evaluation of CCBT for adult depression in terms of clinical implementation and methodological validity, *BMC Psychiatry*, available online at http://www.biomedcentral.com/1471-244X/13/113 Accessed 11th March 2015

Spencer, L.M., and Spencer, S.M. (1993). *Competence at Work*. New York, Wiley.

Spitzer, R.L., Kroenke, K., Williams, J.B., and Lowe, B. (2006). A brief measure for assessing Generalised Anxiety Disorder: the GAD-7. *Archives of Internal Medicine*, 166, 1092-1097.

Steen, S. (2019). A critical appraisal of the economic model underpinning the Improving Access to Psychological Therapies (IAPT) programme. In Jackson, C., and Rizq, R. (Eds) *The Industrialisation of Care: Counselling, Psychotherapy and the Impact of IAPT.* Monmouth, PCCS Books.

Stein, D.M., and Lambert, M.J. (1984). On the relationship between therapist experience and psychotherapy outcome. *Clinical Psychology Review*, 4, 127-142.

Stein, D.M., and Lambert, M.J. (1995). Graduate training in psychotherapy: Are therapy outcomes enhanced? *Journal of Consulting and Clinical Psychology*, 63, 182-196.

Steketee, G.S. (1993). *Treatment of Obsessive Compulsive Disorder .* New York:Guilford Press

Stewart, C.N. (2011). *Research Ethics for Scientists: A Companion for Students*. Chichester, West Sussex, John Wiley and Sons.

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

Stott, R., Wild, J., Grey, N., Liness, S., Warnock-Parkes, E., Commins, S., … Clark, D. M. (2013). Internet-Delivered Cognitive Therapy for Social Anxiety Disorder: A Development Pilot Series. *Behavioural and Cognitive Psychotherapy*, *41* (1), 383-397.

Strathdee, G. (2013) *Keynote speech and launch of the NIHR Mindtech HTC,* available online at:
http://mindtech.org.uk/mindtech-annual-conference/59-exhibitors-at-official-launch-11th-november-2013.html
Accessed 3rd April 2019.

Strauss, B.M., Lutz, W., Steffanowski, A., Wittmann, W.W., Boehnke, J.R., Rubel, J., Scheidt, C.E., Caspar, F., Vogel, H., Altmann, U., Steyer, R., Zimmermann, A, Bruckmayer,E., von Heymann, F, Kramer, D., and Kirchmann, H. (2015). Benefits and challenges in practice-oriented psychotherapy research in Germany: The TK and Qs-PSY-BAY projects of quality assurance in outpatient psychotherapy. *Psychotherapy Research*, 25, 32-51.

Strunk, D.R., Brotman, M.A., DeRubeis, R.J., and Hollon, S.D. (2010). Therapist competence in cognitive therapy for depression: predicting subsequent symptom change. *Journal of Consulting and Clinical Psychology*, 78, 429-437.

Suler, J. (2015). *Psychology of the Digital Age: Humans Become Electric,* Cambridge, Cambridge University Press.

Szentagotai, A. and David, D. (2010) The efficacy of cognitive-behavioral therapy in bipolar disorder: A quantitative meta-analysis. *Journal of Clinical Psychiatry*, 1 (2), 66-72.

Teasdale, J.D., Segal, Z.V., Williams, J.M., Ridgeway, V.A., Soulsby, J.M., and Lau, M.A. (2010). Prevention of relapse/recurrence in major depression by mindfulness-based cognitive therapy. *Journal of Consulting and Clinical Psychology*, 68, 615-623.

The Guardian (2015). Mental health workers protest at move to integrate clinic with job centre. The Guardian 26th June, 2015. Available online at https://www.theguardian.com/society/2015/jun/26/mental-health-protest-clinic-jobcentre-streatham
Accessed 25th October 2019

The Guardian (2016). Depression kills so why aren't people getting better NHS care. The Guardian, 16th March, 2016. Available online at: https://www.theguardian.com/healthcare-network/2016/mar/16/depression-suicide-nhs-mental-health-problem#comments
Accessed 14th September 2019

Thompson, A.G., and France, E.F. (2010). One stop or full stop? The continuing challenges for researchers despite the new streamlined NHS research governance process. (Research article, National Health Service Report), *BMC Health Services Research*, available online at http://www.biomedcentral.com/1472-6963/10/124
Accessed 15th May 2019.

Tracey, T.J., Wampold, B.E., Lichtenberg, J.W., and Goodyear, R.K. (2015). Expertise in Psychotherapy: An elusive goal? *American Psychologist,* 69, 218-229.

Trepka, C., Rees, A., Shapiro, D.A., Hardy, G.E., and Barkham, M. (2004). Therapist competence and outcome of cognitive therapy for depression. *Cognitive Therapy and Research,* 28, 143-157.

Turpin, G., and Wheeler, S. (2011). *IAPT Supervision Guidance*. Available online at https://www.uea.ac.uk/documents/246046/11919343/IAPT+Supervision+Guidance+2011.pdf/3a8f6c76-cea0-4f76-af78-e2f0a1f64c5e
Accessed 24th October 2019

Vallis,T.M., Shaw, B.F., and Dobson, K.S. (1986). The Cognitive Therapy Scale: Psychometric properties. *Journal of Consulting and Clinical Psychology*, 54, 381-385.

van Stolk, C., Hofman, J., Hafner, M., and Janta, B. (2014). *Psychological Wellbeing and Work: Improving Service Provision and Outcomes.* Available online at https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/273433/psychological-wellbeing-and-work.pdf
Accessed 9th October 2019.

Veale, D. (2007). Cognitive Therapy for Obsessive Compulsive Disorder. Advances in Psychiatric Treatment. Available online at:
http://apt.rcpsych.or/content/13/6/438.full Accessed on 10th June 2019.

Voinescu, B.I., Szentagotai, A. and David, D. (2013). Internet administered cognitive-behaviour therapy for insomnia. *Journal of Cognitive and Behavioural Psychotherapies*, 13 (1) 225-237.

Volante, L. (2004). Teaching to the test: What every educator and policy maker should know. Canadian Journal of Educational Administration and Policy. Available online at https://files.eric.ed.gov/fulltext/EJ848235.pdf
Accessed 26th October 2019.

Walfish, S., McAlister, B., O'Donnell, P., & Lambert, M. J. (2012). An investigation of self-assessment bias in mental health providers. *Psychological Reports, 110*, 639–644.

Waller, G. (2009). Evidence-based treatment and therapist drift. *Behaviour Research and Therapy*, 47, 119-127.

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

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

Walsh, P., Carroll, J., and Sleator, R.D. (2013). Accelerating in silico research with workflows: A lesson in simplicity. *Computers in Biology and Medicine*, 43, 2028-2035.

Watson, J.B. (1913). Psychology as the behaviorist views It. *Psychological Review*, 20, 158-177.

Wampold, B.E. (2001) *The Great Psychotherapy Debate: Models, Methods and Findings*. Mahwah, New Jersey, Routledge.

Wampold, B.E. (2015). How important are the common factors in psychotherapy? An update. *World Psychiatry*, 14, 270-277.

Wampold, B.E., and Brown, G.S. (2005). Estimating variability in outcomes attributable to therapists? A naturalistic study of outcomes in managed care. *Journal of Consulting and Clinical Psychology*, 73, 914-923.

Wampold, B.E., Baldwin, S.A., Holtforth, M.G., and Imel, Z.E. (2017). What characterises effective therapists? In Castonguay, L.G., and Hill, C.E.,(Eds.). (2017). *How and Why are Some Therapists Better Than Others? Understanding Therapist Effects*. American Psychological Association, Washington DC.

Wampold, B.E., Fluckiger, C., Del Re, A.C., Yulish, N.E., Frost, N.D., Pace, B.T., Goldberg, S.B., Miller, S.D., Baardseth, T.P., Laska, K.M., and Hilsenroth, M.J. (2016).In pursuit of truth> A critical examination of meta-analyses of cognitive behavior therapy. *Psychotherapy Research,* 27, 14-32.

Wampold, B.E., Goodheart, C.D., and Levant, R.F. (2007). Clarification and elaboration on evidence-based practice in psychology. *American Psychologist,* 62, 616-618.

Wampold, B.E., and Imel, Z.E. (2015). *The Great Psychotherapy Debate: The Evidence for What Makes Psychotherapy Work.* New York, Routledge.

Webb, C.A., DeRubeis, R.J., and Barber, J.P. (2010). Therapist adherence/competence and treatment outcome: A meta-analytic review. J*ournal of Consulting and Clinical Psychology*, 78, 200-211.

Weck, F., Richtberg, S., Jakob, M., Neng, J.M.B., and Hofling, V. (2015). Therapist competence and therapeutci allinace are important in the treatment of health anxiety (hypochondriasis). *Psychiatry Research*, 228, 53-58.

Weck, F., Grikscheit, F., Hofling, V., Kordt, A., Hamm, A.O., Gerlach, A.L., Alpers, G.W., Arolt, V., Kircher, T., Pauli, P., Rief, W., and Lang, T. (2016). The role of treatment delivery factors in exposure-based cognitive behavioural therapy of panic disorder with agoraphobia. *Anxiety Disorders*, 42, 10-18.

Weck, F., Kauffmann, Y.M. and Hofling, V. (2017). Competence feedback improves competence in trainee therapists: A randomised controlled pilot study. *Psychotherapy Research*, 27, 501-509.

Weiss, D.S. (2007). The Impact of Event Scale-Revised. In J.P. Wilson, & T.M. Keane (Eds.) *Assessing Psychological Trauma and PTSD: a Practitioner's Handbook* (2nd ed., pp. 168-189). New York: Guilford Press.

Wells, A. (1997). Cognitive therapy of anxiety disorders: A practice manual and conceptual guide. Wiley, Chichester.

We Need to Talk Coalition Report. (2010). We Need to Talk: Getting the Right Therapy at the Right Time. Available online at https://www.mind.org.uk/media/280583/We-Need-to-Talk-getting-the-right-therapy-at-the-right-time.pdf
Accessed 26th October 2019.

Westbrook, D. and Kirk, J. (2005) The clinical effectiveness of cognitive behaviour therapy: Outcome for a large sample of adults treated in routine practice, *Behaviour Research and Therapy,* 43 (6), 1243-1261.

Westwood, S., Morison, L., Allt, J., and Holmes, N. (2017). Predictors of emotional exhaustion, disengagement and burnout among improving access to psychological therapies (IAPT) practitioners. J*ournal of Mental Health,* 26, 172-179.

Whisman, M. (1993). Mediators and moderators of change in cognitive therapy of depression. *Psychological Bulletin*, 114, 248-265.

Williams, C.H. (2015). Improving access to psychological therapies (IAPT) and treatment outcomes: Epistemological assumptions and controversies. *Journal of Psychiatric and Mental Health Nursing*, 22, 344-351.

Williams, B., Boyle, M., and O'Meara, P. (2009). Can undergraduate paramedic students accurately identify lung sounds? *Emergency Medicine*, 26, 580-582.

Williams, R.M., Moorey, S., and Cobb, J. (1991). training in cognitive-behaviour therapy: Pilot evaluation of a training course using the Cognitive Therapy Scale. *Behavioural and Cognitive Psychotherapy*, 19, 373-376.

Wolf, A.W., and Goldfried, M.R. (2014). Clinical experiences in using cognitive-behaviour therapy to treat panic disorder. *Behaviour Therapy*, 45, 36-46.

World Health Organisation (2008). *The Global burden of disease: 2004 update*. Available online at
https://www.who.int/healthinfo/global_burden_disease/GBD_report_2004update_full.pdf?ua=1
Accessed 25th October 2019

World Health Organisation (2011). *Mental Health Atlas 2011*. Available online at https://apps.who.int/iris/bitstream/handle/10665/44697/9799241564359_eng.pdf;jsessionid=F94FB3498E96ABBDE61BB5017640FD31?sequence=1
Accessed 25th October 2019

Worthen, V., and Lambert, M.J. (2007). Outcome orientated supervision: Advantages of adding systematic client tracking to supportive consultations. *Counselling and Psychotherapy Research*, 7, 48-53.

Youn, S.J., Xiao, H., Kim, H., Castonguay, L.G., McAleavey, A.A., Newman, M.G., and Safran, J.D., (2017) Effective and less effective therapists for Generalised Anxiet Disorder: Are they conducting therapy the same way? In Castonguay, L.G., and Hill, C.E.,(Eds.). (2017). *How and Why are Some Therapists Better Than Others? Understanding Therapist Effects*. American Psychological Association, Washington DC.

Young, J. (1999). *Cognitive Therapy for Personality Disorders: A Schema-Focussed Approach.* Sarasota, Florida, Professional Resource Press.
Young, J., and Beck A.T. (1980). *Cognitive Therapy Scale Rating Manual.* Unpublished manuscript, Beck Institute, Philadelphia ,Pennsylvania.

Young, J., and Beck, A.T. (1988). *Cognitive Therapy Scale.* Unpublished manuscript, University of Pennsylvania, Philadelphia, Pennsylvania.

Zarafontis-Muller, S., Kuhr, K., and Bechdolf, A. (2014). The relationship between therapist's competence and adherence to outcome in cognitive-behavioural therapy-results of a meta-analysis. *Fortschritte der Neurologie-Psychiatrie*, 82, 502-510.

Zinbarg, R.E. Craske , M.G. & Barlow D.A. (2006) . *Mastery of your anxiety and worry.* Oxford, Oxford University Press.

Zou, J.B., Dear, B.F., Titov, N., Lorian, C.N., Johnston, L., Spence, J., Knight, R.G., Anderson, T., and Sachdev, P. (2012). Brief internet-delivered cognitive-behavioral therapy for anxiety in older adults: A feasibility trial. *Journal of Anxiety Disorders*, 26, 650-655.

APPENDICES

Appendix 1 Ewbank, Cummins, Tablan, Bateup, Catarino, Martin and Blackwell (2019)

JAMA Psychiatry | Original Investigation

Quantifying the Association Between Psychotherapy Content and Clinical Outcomes Using Deep Learning

Michael P. Ewbank, PhD; Ronan Cummins, PhD; Valentin Tablan, PhD; Sarah Bateup, MA; Ana Catarino, PhD; Alan J. Martin, BSc; Andrew D. Blackwell, PhD

Supplemental content

IMPORTANCE Compared with the treatment of physical conditions, the quality of care of mental health disorders remains poor and the rate of improvement in treatment is slow, a primary reason being the lack of objective and systematic methods for measuring the delivery of psychotherapy.

OBJECTIVE To use a deep learning model applied to a large-scale clinical data set of cognitive behavioral therapy (CBT) session transcripts to generate a quantifiable measure of treatment delivered and to determine the association between the quantity of each aspect of therapy delivered and clinical outcomes.

DESIGN, SETTING, AND PARTICIPANTS All data were obtained from patients receiving Internet-enabled CBT for the treatment of a mental health disorder between June 2012 and March 2018 in England. Cognitive behavioral therapy was delivered in a secure online therapy room via instant synchronous messaging. The initial sample comprised a total of 17 572 patients (90 934 therapy session transcripts). Patients self-referred or were referred by a primary health care worker directly to the service.

EXPOSURES All patients received National Institute for Heath and Care Excellence-approved disorder-specific CBT treatment protocols delivered by a qualified CBT therapist.

MAIN OUTCOMES AND MEASURES Clinical outcomes were measured in terms of reliable improvement in patient symptoms and treatment engagement. Reliable improvement was calculated based on 2 severity measures: Patient Health Questionnaire (PHQ-9) and Generalized Anxiety Disorder 7-item scale (GAD-7), corresponding to depressive and anxiety symptoms respectively, completed by the patient at initial assessment and before every therapy session.

RESULTS Treatment sessions from a total of 14 899 patients (10 882 women) aged between 18 and 94 years (median age, 34.8 years) were included in the final analysis. We trained a deep learning model to automatically categorize therapist utterances into 1 or more of 24 feature categories. The trained model was applied to our data set to obtain quantifiable measures of each feature of treatment delivered. A logistic regression revealed that increased quantities of a number of session features, including change methods (cognitive and behavioral techniques used in CBT), were associated with greater odds of reliable improvement in patient symptoms (odds ratio, 1.11; 95% CI, 1.06-1.17) and patient engagement (odds ratio, 1.20, 95% CI = 1.12-1.27). The quantity of nontherapy-related content was associated with reduced odds of symptom improvement (odds ratio, 0.89; 95% CI, 0.85-0.92) and patient engagement (odds ratio, 0.88, 95% CI, 0.84-0.92).

CONCLUSIONS AND RELEVANCE This work demonstrates an association between clinical outcomes in psychotherapy and the content of therapist utterances. These findings support the principle that CBT change methods help produce improvements in patients' presenting symptoms. The application of deep learning to large clinical data sets can provide valuable insights into psychotherapy, informing the development of new treatments and helping standardize clinical practice.

JAMA Psychiatry. doi:10.1001/jamapsychiatry.2019.2664
Published online August 22, 2019.

Author Affiliations: Clinical Science Laboratory, Ieso Digital Health, Cambridge, England.

Corresponding Author: Michael P. Ewbank, PhD, Clinical Science Laboratory, Ieso Digital Health, Cowley Road, The Jeffreys Building, Milton, Cambridge CB4 0DS, England (m.ewbank@iesohealth.com).

Compared with treatment of physical conditions, the quality of care of mental health disorders remains poor, and the rate of improvement in treatment is slow.[1] Outcomes for many mental disorders have stagnated or even declined since the original treatments were developed.[2,3] A primary reason for the gap in quality of care is the lack of systematic methods for measuring the delivery of psychotherapy.[1] As with any evidence-based intervention, to be effective, treatment needs to be delivered as intended (also known as treatment integrity),[4,5] which requires accurate measurement of treatment delivered.[6] However, while it is relatively simple to monitor the delivery of most medical treatments (eg, the dosage of a prescribed drug), psychotherapeutic treatments are a series of private discussions between the patient and clinician. As such, monitoring the delivery of this type of treatment to the same extent as physical medicine would require infrastructure and resources beyond the scope of most health care systems.

The National Institute for Heath and Care Excellence and the American Psychological Association recommend cognitive behavioral therapy (CBT) as a treatment for most common mental health problems such as depression and anxiety-related disorders. Cognitive behavioral therapy refers to a class of psychotherapeutic interventions informed by the principle that mental disorders are maintained by cognitive and behavioral phenomena and that modifying these maintaining factors helps produce enduring improvements in patients' presenting symptoms.[7,8] Despite its widespread use, the Improving Access to Psychological Therapies (IAPT) program in England includes no objective measure of treatment integrity for CBT, and it has been proposed that only 3.5% of psychotherapy randomized clinical trials use adequate treatment integrity procedures.[9]

Understanding how CBT works is of particular interest given that the relative effects of different psychotherapeutic interventions appear similar.[10] Thus, whether treatments work through specific factors (eg, CBT change methods) or factors common to most psychotherapies (eg, therapeutic alliance) remains a core issue in the field.[11,12] Studies commonly use observational coding methods (eg, ratings/transcription of recorded therapeutic conversations) to investigate the association between treatment delivered and outcomes.[5] Owing to the resource-intensive nature of this method, studies typically focus on a small number of therapeutic components in a relatively small sample of patients. As with many randomized clinical trials, the results of such interventions are difficult to transfer to real-world psychotherapy[13] and require sample sizes larger than typically used.[14] To determine the most effective components of CBT and whether CBT works via the mechanisms proposed by the approach,[15] quantifiable measures of treatment delivered need to be obtained in a natural clinical context and be gathered from a sufficiently large enough sample to draw meaningful conclusions.

Here, we used a large-scale data set containing session transcripts from more than 14 000 patients receiving internet-enabled CBT (IECBT) (approximately 90 000 hours of therapy). In IECBT, a patient communicates with a qualified CBT therapist using a real-time text-based message system. Internet-enabled CBT has been shown to be clinically effective for the treatment of depression[16] and is currently deployed within IAPT. Using a deep learning approach, we developed a model to automatically categorize therapist utterances according to the role that they play in therapy, generating a quantifiable measure of treatment delivered. We then investigated the association between the quantity of each aspect of therapy delivered and clinical outcomes.

Key Points

Question What aspects of psychotherapy content are significantly associated with clinical outcomes?

Findings In this quality improvement study, a deep learning model was trained to automatically categorize therapist utterances from approximately 90 000 hours of internet-enabled cognitive behavior therapy (CBT). Increased quantities of CBT change methods were positively associated with reliable improvement in patient symptoms, and the quantity of nontherapy-related content showed a negative association.

Meaning The findings support the key principles underlying CBT as a treatment and demonstrate that applying deep learning to large clinical data sets can provide valuable insights into the effectiveness of psychotherapy.

Methods

Design

Data were obtained from patients receiving IECBT for the treatment of a mental health disorder between June 2012 and March 2018. Internet-enabled CBT was delivered using a commercial package currently used in the English National Health Service, provided by Ieso Digital Health (https://www.iesohealth.com/), following internationally recognized standards for information security (ISO 27001; https://www.iesohealth.com/en-gb/legal/iso-certificates). The National Institute for Heath and Care Excellence approved disorder-specific CBT treatment protocols,[17] based on Roth and Pilling CBT competences framework,[18] were delivered in a secure online therapy room via instant synchronous messaging by a British Association for Behavioral and Cognitive Psychotherapies–accredited CBT therapist (see eFigure 1 in the Supplement for a realistic example of a therapy conversation). Patients self-referred or were referred by a primary health care worker directly to the service.

The IAPT program is a large-scale initiative aimed at increasing access to evidence-based psychological therapy for common mental health disorders within the English National Health Service.[19] The information captured through IAPT's minimum data set is intended to support monitoring of implementation and effectiveness of national policy/legislation, performance analysis and benchmarking, and national audit of IAPT services. As determined by the National Health Service, and per The National Institute of Health and Care Excellence principles,[20] clinical audit studies within the IAPT framework do not require additional patient consent or ethical

Box. Feature Categories Used in Transcript Annotation

Therapy Feature Categories

- Hello
- Mood check
- Obtain update
- Bridge
- Risk check
- Set agenda
- Review homework
- Set goals
- Formulation
- Change methods
- Perceptions of change
- Setting homework
- Planning for the future
- Elicit feedback
- Summarize session
- Give feedback
- Arrange next session
- Goodbye
- Socratic questioning[a]
- Therapeutic thanks[a]
- Therapeutic empathy[a]
- Therapeutic praise[a]
- Collaboration[a]
- Other

[a] Features tagged using regular expressions.

approval.[20] When registering to use the Ieso service, patients provide written informed consent as part of a privacy policy agreement, allowing the service to use their anonymized data for audit purposes and to support research, including academic publications.

Clinical Outcomes

Clinical outcomes were defined according to IAPT guidelines[19] and were measured in terms of reliable improvement and IAPT engagement and included as binary measures (ie, 0 or 1). A patient was classed as engaged if they attended 2 or more treatment sessions. Reliable improvement was calculated based on 2 severity measures: Patient Health Questionnaire (PHQ-9)[21] and Generalized Anxiety Disorder 7-item scale (GAD-7),[22] corresponding to depressive and anxiety symptoms respectively, completed by the patient at initial assessment and before every therapy session (see eMethods in the Supplement for details).

Therapy Feature Categories

We defined a total of 24 feature categories (**Box**), informed by the CBT competences framework[18] and the Revised Cognitive Therapy Scale.[23] A research psychologist (M.P.E.) annotated 290 therapy session transcripts, under the guidance of a qualified clinical therapist (S.B.), tagging each therapist text-message utterance as belonging to 1 (or more) of 19 features, with 5 features tagged using regular expressions (see eTable 1 in the Supplement for a full description). A deep learning model (see eMethods in the Supplement) was trained on the annotated utterances and then used to automatically classify all utterances in the full data set into 1 or more of 24 feature categories. Model accuracy is detailed in eTable 2 in the Supplement. To obtain a measure of interrater agreement, a second psychologist (S.B.) annotated a subsample of the transcripts. The interrater reliability was $\kappa = 0.54$ (a value of 0.4-0.6 is considered moderate agreement, with zero equaling chance agreement[24]; eTable 3 in the Supplement).

Statistical Analysis

Using the output of the model, the mean number of words for each feature, averaged across all sessions, was calculated for each case. The final treatment session was excluded because outcome measures are taken prior to the commencement of each treatment session. The initial sample comprised a total of 90 934 session transcripts taken from 17 572 patients, with a reliable improvement rate of 63.4% and IAPT engagement rate of 87.3%.

All analyses were performed in R (the R Foundation). Cases with missing start or end PHQ-9 or GAD-7 scores (n = 1338) were excluded from the analysis. We performed 3 multivariable logistic regression analyses. First, a multivariable logistic regression was performed to investigate the association between session features and reliable improvement. Predictor variables were the mean number of words for each feature across sessions plus patient demographics: starting PHQ-9 and GAD-7 scores, sex (male, female, or unstated/unknown), age, whether the patient had a long-term physical condition (yes, no, or unstated/unknown), and whether the patient was taking psychotropic medication at the start of treatment (prescribed not taking, prescribed taking, not prescribed, or unstated/unknown). The number of sessions completed and the mean duration of sessions were also included. Cases with a mean of fewer than 50 patient words were excluded (n = 16), leaving a total of 13 073 patients (at a clinical caseness threshold and engaged in treatment) in the analysis.

We also investigated the association between first-session features and IAPT engagement. Predictor variables were the number of each therapy feature in the first session, patient demographics, and duration of first session. Sessions with a total of fewer than 50 patient words were excluded (n = 121) making a total of 14 899 patients, at caseness.

Details of a logistic regression analysis investigating the association between first-session features and outcomes can be found in eResults and eTable 5 in the Supplement. Details of diagnoses for patients included in the analysis can be found in eTable 4 in the Supplement. Patient demographic information is shown in **Tables 1** and **2** and eTable 5 in the Supplement.

For all analyses, continuous predictor variables were scaled and centered to the mean. Statistical significance was defined as P less than .05 two-tailed, uncorrected. Multicollinearity analyses revealed that variance inflation factors were smaller than 2 for all predictor variables, confirming that regression models were not affected by the presence of multicollinearity.

Table 1. Factors Associated With Reliable Improvement–All Sessions[a]

Feature	No. of Words, Mean (SD)	Sessions, %	Odds Ratio (95% CI)	z Value	P Value
Hello	12 (22.7)	99.6	0.92 (0.88-0.96)	−3.57	<.001
Mood check	5.6 (7)	97.9	0.99 (0.95-1.03)	−0.34	.73
Obtain update	16.4 (14.5)	59.0	1.03 (0.99-1.08)	1.56	.12
Bridge	12.2 (17.9)	27.9	0.95 (0.91-0.98)	−2.76	.006
Risk check	13.6 (31.5)	21.0	0.85 (0.81-0.89)	−7.54	<.001
Set agenda	47.2 (43.5)	71.3	1.08 (1.02-1.14)	3.02	.002
Review homework	18.5 (19.2)	44.5	1.04 (1.00-1.09)	2.00	.04
Set goals	15.9 (30.8)	19.4	1.00 (0.96-1.05)	0.40	.69
Formulation	30.3 (63.9)	18.2	0.96 (0.92-1.00)	−1.89	.06
Give feedback	33.6 (40)	52.1	1.05 (1.00-1.10)	2.20	.02
Change methods	477.1 (236)	97.9	1.11 (1.06-1.17)	4.37	<.001
Perceptions of change	1.6 (4.8)	5.8	1.11 (1.06-1.16)	4.59	<.001
Set homework	63.2 (48.9)	69.1	0.96 (0.92-1.00)	−1.68	.09
Planning for future	1.1 (6)	2.4	1.12 (1.06-1.19)	4.01	<.001
Elicit feedback	15.3 (16.4)	55.3	1.06 (1.02-1.11)	2.82	.004
Summarize session	0.25 (2.6)	0.4	0.99 (0.95-1.03)	−0.52	.60
Arrange next session	30 (21.3)	82.5	1.00 (0.96-1.04)	0.05	.96
Goodbye	15.4 (10.4)	90.7	0.95 (0.91-0.99)	−2.34	.02
Socratic questioning	24.1 (31.1)	47.4	1.02 (0.98-1.06)	0.95	.34
Therapeutic thanks	5.4 (13.3)	13.3	0.97 (0.93-1.01)	−1.48	.14
Therapeutic empathy	21 (31.3)	38.0	0.84 (0.81-0.88)	−8.21	<.001
Therapeutic praise	30.6 (39.4)	52.6	1.21 (1.15-1.27)	7.18	<.001
Collaboration	41 (45.9)	61.9	0.97 (0.93-1.02)	−1.09	.27
Other	121.1 (81)	96.0	0.88 (0.85-0.92)	−5.82	<.001
Variable, mean/prevalence (SD)					
Total sessions, No.	6.2 (2.9)	NA	1.22 (1.17-1.27)	9.01	<.001
Session duration, min	62.4 (7.5)	NA	0.95 (0.91-0.99)	−2.34	.02
Start PHQ-9	14.7 (5.4)	NA	0.95 (0.91-0.99)	−2.41	.03
Start GAD-7	8.3 (5.7)	NA	1.29 (1.23-1.34)	11.8	<.001
Patient age, y	34.8 (12.0)	NA	1.16 (1.12-1.22)	7.47	<.001
Patient sex, No. (%)					
Male	3493 (26.7)	NA	0.96 (0.88-1.05)	−0.89	.50
Female	9537 (72.9)	NA			
Unknown/not stated	43 (0.4)	NA	0.92 (0.49-1.78)	−0.24	.74
Long-term condition, No. (%)					
No	6056 (46.4)	NA			
Yes	3632 (27.8)	NA	0.72 (0.66-0.80)	−6.55	<.001
Unknown/not stated	3383 (25.8)	NA	0.78 (0.71-0.86)	−5.08	<.001
Psychotropic medication, No. (%)					
Prescribed not taking	1116 (8.6)	NA			
Not prescribed	5971 (45.7)	NA	1.23 (1.06-1.41)	2.84	.004
Prescribed taking	5535 (42.3)	NA	0.98 (0.84-1.13)	−0.27	.78
Unknown/not stated	451 (3.4)	NA	0.85 (0.67-1.08)	−1.28	.20

Abbreviations: GAD-7, Generalized Anxiety Disorder 7-Item scale; NA, not applicable; PHQ-9, Patient Health Questionnaire.

[a] Output of logistic regression investigating association between reliable improvement and mean number of words per feature across treatment. Standardized odds ratios indicate the association of an increase of 1 SD of a feature with the odds of improvement. Percentage of sessions indicates the percentage of the total number of sessions that contained utterances categorized as that feature. Female sex, no long-term conditions, and prescribed not taking psychotropic medication were reference classes for the categorical variables.

Results

Factors Associated With Reliable Improvement Across Treatment

Figure 1 shows the standardized odds ratios (ORs) for each therapy feature included in the multivariable logistic regression (Table 1). The results revealed increased quantities of "therapeutic praise" (OR, 1.21; 95% CI, 1.15-1.27), "planning for the future" (OR, 1.12; 95% CI, 1.06-1.19), "perceptions of change" (OR, 1.11; 95% CI, 1.06-1.16), "change methods" (OR, 1.11; 95% CI, 1.06-1.17), "set agenda" (OR, 1.08; 95% CI, 1.02-1.14), "elicit feedback" (OR, 1.06; 95% CI, 1.02-1.11), "give feedback" (OR, 1.05; 95% CI, 1.00-1.10), and "review homework" (OR, 1.04; 95% CI, 1.00-1.09) were all associated with greater odds of reliable improvement. By contrast, increases in nontherapy-related content ("other" [OR, 0.89; 95% CI, 0.85-0.92], "hello" [OR, 0.92; 95% CI, 0.88-0.96], and "goodbye" [OR, 0.95; 95% CI, 0.91-

Table 2. First-Session Factors Associated With IAPT Engagement[a]

Feature	No. of Words, Mean (SD)	Sessions, %	Odds Ratio (95% CI)	z Value	P Value
Hello	14.4 (34.7)	99.7	0.93 (0.88-0.99)	−2.45	.01
Mood check	5.6 (10.5)	48.1	0.98 (0.93-1.03)	−0.96	.33
Obtain update	12.3 (19.6)	46.1	0.96 (0.92-1.01)	−1.54	.11
Bridge	9.7 (24.8)	22.7	0.94 (0.90-0.98)	−2.63	.008
Risk check	22.8 (54.7)	30.4	0.98 (0.94-1.03)	−0.69	.48
Set agenda	61.3 (68.7)	74.9	0.99 (0.94-1.05)	−0.27	.79
Review homework	15.2 (27.3)	39.4	0.96 (0.91-1.01)	−1.47	.14
Set goals	28.3 (57.9)	35.9	1.03 (0.98-1.09)	1.07	.28
Formulation	53.2 (126)	30.4	1.10 (1.04-1.17)	3.33	<.001
Give feedback	17.4 (57.2)	49.3	1.00 (0.95-1.07)	0.31	.75
Change methods	426.5 (279.5)	97.6	1.20 (1.12-1.27)	5.56	<.001
Perceptions of change	1.13 (7.4)	3.6	0.97 (0.93-1.01)	−1.42	.14
Set homework	75.8 (74.4)	78.4	1.09 (1.03-1.16)	2.97	<.002
Planning for future	0.56 (8.5)	1.0	0.93 (0.89-0.96)	−3.77	<.001
Elicit feedback	17.4 (25)	60.9	1.09 (1.03-1.16)	2.97	.002
Summarize session	0.24 (4.67)	0.3	1.00 (0.94-1.09)	0.01	.98
Arrange next session	33.1 (32.6)	84.0	1.17 (1.10-1.24)	5.30	<.001
Goodbye	16.2 (15.6)	90.9	1.02 (0.97-1.08)	0.83	.40
Socratic questioning	20 (39.5)	40.8	0.94 (0.89-0.99)	−2.28	.02
Therapeutic thanks	8.5 (24.3)	19.4	1.13 (1.06-1.20)	3.73	<.001
Therapeutic empathy	25.5 (51.1)	44.0	0.93 (0.88-0.97)	−3.20	.001
Therapeutic praise	23.3 (47)	41.8	1.05 (0.98-1.11)	1.47	.15
Collaboration	45.2 (72.8)	60.4	1.01 (0.94-1.07)	0.26	.79
Other	141.1 (117.4)	96.9	0.88 (0.84-0.92)	−5.12	<.001
Variable, mean/prevalence (SD)					
Session duration, min	63.1 (9.9)	NA	1.26 (1.20-1.33)	8.89	<.001
Start PHQ-9	14.9 (5.5)	NA	0.87 (0.82-0.92)	−4.81	<.001
Start GAD-7	8.8 (5.9)	NA	1.00 (0.95-1.06)	−0.01	.99
Patient age	34.8 (12.0)	NA	1.07 (1.02-1.13)	2.64	.008
Patient sex, %					
Male	3967 (26.7)	NA	1.02 (0.91-1.01)	0.28	.78
Female	10 882 (73.0)	NA	NA	NA	NA
Unknown/not stated	50 (0.3)	NA	0.95 (0.45-2.34)	−0.11	.91
Long-term condition, %					
No	6860 (46.0)	NA			
Yes	4129 (27.7)	NA	1.02 (0.90-1.15)	0.24	.81
Unknown/not stated	3910 (26.3)	NA	0.90 (0.80-1.02)	−1.68	.09
Psychotropic medication, %					
Prescribed not taking	1304 (8.8)				
Not prescribed	6755 (45.3)	NA	1.21 (1.02-1.44)	2.19	.03
Prescribed taking	6320 (42.4)	NA	1.20 (1.01-1.47)	2.06	.04
Unknown/not stated	520 (3.5)	NA	1.10 (1.01-1.43)	0.64	.52

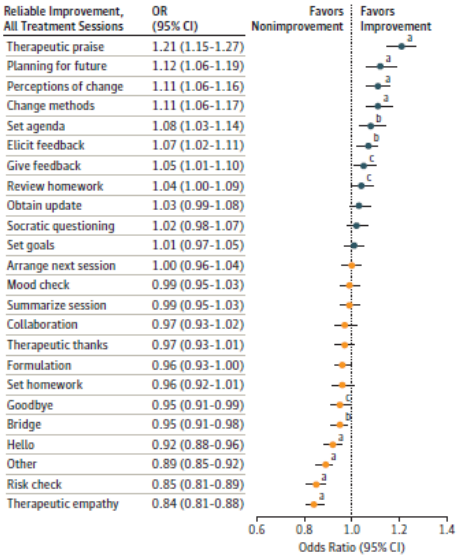
Abbreviations: GAD-7, Generalized Anxiety Disorder 7-item scale; IAPT, Improving Access to Psychological Therapies; NA, not applicable; PHQ-9, Patient Health Questionnaire.

[a] Output of logistic regression investigating association between patient engagement and number of words per feature in the first treatment session. Standardized odds ratios indicate the effect of an increase of 1 SD of a feature on the odds of engagement. Percentage of sessions indicates the percentage of the total number of first treatment sessions that contained utterances categorized as that feature. Female sex, no long-term conditions, and prescribed not taking psychotropic medication were reference classes for the categorical variables.

0.99]), along with "therapeutic empathy" (OR, 0.84; 95% CI, 0.81-0.88), "risk check" (OR, 0.85; 95% CI, 0.81-0.89), and "bridge" (OR, 0.95; 95% CI, 0.91-0.98) were negatively associated with improvement.

Patient variables of starting GAD-7 score (OR, 1.29; 95% CI, 1.23-1.34), not being prescribed medication (OR, 1.23; 95% CI, 1.06-1.41), patient age (OR, 1.16; 95% CI, 1.12-1.22), and total number of treatment sessions (OR, 1.22; 95% CI, 1.17-1.27) were also associated with increased odds of improvement. Starting PHQ-9 score (OR, 0.95; 95% CI, 0.91-0.99), the presence of a long-term medical condition (OR, 0.72; 95% CI, 0.66-0.88), and longer session durations (OR,

Figure 1. Factors Associated With Reliable Improvement-All Sessions



Forest plot of logistic regression model investigating association between mean number of words per feature across treatment and reliable improvement. Standardized odds ratios and 95% confidence intervals are shown (and listed in the right column). Adjusted for total number of sessions, symptom severity, patient sex, age, medication status, presence of long-term condition, and session duration.

[a] $P < .001$.

[b] $P < .01$.

[c] $P < .05$.

0.95; 95% CI, 0.91-0.99) were associated with reduced odds of improvement.

Factors Associated With IAPT Engagement in First Treatment Session

Figure 2 shows the standardized ORs for each session feature included in the multivariable logistic regression (Table 2). We found that "change methods" (OR, 1.20; 95% CI. 1.12-1.27), "elicit feedback" (OR, 1.09; 95% CI. 1.03-1.16), "set homework" (OR, 1.09; 95% CI. 1.03-1.16)," arrange next session" (OR, 1.17; 95% CI. 1.10-1.24), "therapeutic thanks" (OR, 1.13; 95% CI. 1.06-1.20), and "formulation" (OR, 1.10; 95% CI, 1.04-1.17) were associated with increased odds of IAPT engagement. By contrast, nontherapy-related content ("other" and "hello") showed a negative association ("other" OR, 0.88; 95% CI, 0.84-0.92; "hello" OR, 0.93; 95% CI, 0.88-0.99), as did "therapeutic empathy" (OR, 0.93; 95% CI, 0.88-0.97), "Socratic questioning" (OR, 0.94; 95% CI, 0.89-0.99), "bridge" (OR, 0.94; 95% CI, 0.90-0.98), and "planning for the future" (OR, 0.93; 95% CI, 0.89-0.96). Patient age (OR, 1.07; CI, 1.02-1.13), not being prescribed medication (OR, 1.21; 95% CI, 1.02-1.44), being prescribed and taking medication (OR, 1.20; 95% CI, 1.01-1.47), and duration of the first session (OR, 1.26; CI, 1.20-1.33) were positively associated with IAPT engagement, while starting PHQ-9 score (OR, 0.87; CI, 0.82-0.92) was negatively associated.

Discussion

Improving the quality and efficacy of psychotherapy requires that treatment be delivered as intended; however, monitoring and measuring treatment delivered presents a substantial challenge. We developed a method of objectively quantifying psychotherapy using a deep learning approach to automatically categorize therapist utterances from approximately 90 000 hours of IECBT. We find that factors specific to CBT, as well as factors common to most psychotherapies, are associated with increased odds of reliable improvement in patient symptoms.

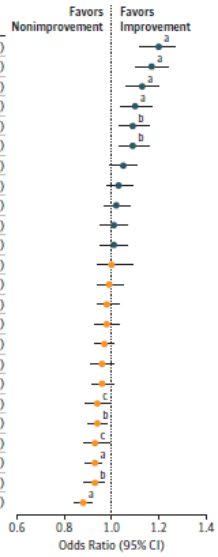
The results revealed a positive association between the quantity of CBT change method-related content and both reliable improvement and IAPT engagement. This finding supports the key principles underlying CBT and provides validation for CBT as a treatment (ie, modifying cognitive and behavioral factors produces improvements in patient symptoms). Here, the category of "change methods" included any example of cognitive or behavioral reattribution, skill-teaching, conceptualization, or psychoeducation. Thus, further research is needed to determine the association between different types of change method and outcomes.[15]

Homework in CBT is used to help patients practice skills learned in therapy and generalize these skills to the real world.[25] Increased content related to reviewing homework was positively associated with symptom improvement, while setting homework in the first session was associated with increased engagement. It is unclear whether an increase in reviewing homework plays a causal role in symptom change or whether it reflects a patient who has completed homework; however, these findings accord with evidence that out-of-session homework is important in determining outcomes in CBT.[26] The results show that agenda setting is also positively associated with reliable improvement. Agenda setting involves the therapist and patient deciding on the topics to be discussed during the session. However, we are unable to determine whether the agenda was adhered to in the session. The results also support the principle that giving and eliciting feedback helps both the therapist and patient develop a greater understanding of key issues and possibly strengthens the therapeutic alliance.[27]

Session content related to planning for the future after therapy and discussing perceptions of change was also positively associated with improvement. A discussion of perceptions of change is only likely to occur following some degree of change; similarly, planning for a future most likely occurs when patients are close to completing treatment and/or have moved toward improvement. As such, the increased occurrence of both features is likely to be reflective of treatment progressing well. Consistent with this, neither feature was sig-

Figure 2. First-Session Factors Associated With IAPT Engagement

IAPT Engagement, First Treatment Sessions	OR (95% CI)
Change methods	1.20 (1.12-1.27)
Arrange next session	1.17 (1.10-1.24)
Therapeutic thanks	1.13 (1.06-1.20)
Formulation	1.10 (1.04-1.17)
Elicit feedback	1.09 (1.03-1.16)
Set homework	1.09 (1.03-1.16)
Therapeutic praise	1.05 (0.99-1.11)
Set goals	1.03 (0.98-1.09)
Goodbye	1.02 (0.97-1.08)
Give feedback	1.01 (0.95-1.07)
Collaboration	1.01 (0.95-1.07)
Summarize session	1.00 (0.94-1.09)
Set agenda	0.99 (0.94-1.05)
Risk check	0.98 (0.94-1.03)
Mood check	0.98 (0.93-1.03)
Perceptions of change	0.97 (0.93-1.01)
Review homework	0.96 (0.91-1.01)
Obtain update	0.96 (0.92-1.01)
Socratic questioning	0.94 (0.89-0.99)
Bridge	0.94 (0.90-0.98)
Hello	0.93 (0.88-0.99)
Planning for future	0.93 (0.89-0.96)
Therapeutic empathy	0.93 (0.88-0.97)
Other	0.88 (0.84-0.92)



Forest plot of logistic regression model investigating association between mean number of words per feature in the first treatment session and patient engagement. Standardized odds ratios and 95% confidence intervals are shown (and listed in the right column). Adjusted for symptom severity, patient sex, age, medication status, presence of long-term condition, and session duration.

[a] $P < .001$.

[b] $P < .01$.

[c] $P < .05$.

nificantly associated with outcomes in the first treatment session (eTable 5 in the Supplement). By contrast, goal setting in the first session was positively associated with improvement, supporting the goal-directed nature of CBT.[27] Content associated with formulation (ie, the beliefs and behavioral strategies that characterize a disorder)[28] in the first session also showed a positive association with IAPT engagement (and a borderline significant association with improvement), suggesting that placing patients' experiences within a cognitive behavioral framework early in therapy is beneficial.

Several features were found to be negatively associated with outcomes, in particular nontherapy-related content. Content that did not fall within any of the other 23 categories ("other") includes utterances related to technical/practical matters or nontherapeutic advice/conversations. While greetings and goodbyes are essential to the structure of a therapy session, our results indicate that, when aggregated across sessions, an excessive or disproportionate amount of time spent on such nontherapeutic aspects may reduce the quantity of active intervention. Importantly, this suggests that rather than the quantity of conversation, it is the therapeutic nature of conversation and/or the dosage of therapy delivered in a session that is associated with improvement in patient symptoms.

Risk checking also showed a strong negative association with reliable improvement. We believe this is likely to be reflective of patients with more complex problems who report more thoughts of self-harm. The quantity of risk checks will increase if a patient confirms that they feel at risk; thus, it is important to recognize that increased risk-checking content is essential and unavoidable. An extended period focused on risk is also likely to cause a deviation in the structure of the session and a subsequent reduction in the dosage of active therapy delivered.

A central issue in psychotherapy research is whether different approaches work through specific factors or factors that are common to most psychotherapies. Here, we find a positive association between improvement and/or IAPT engagement for each of 6 techniques identified as distinguishing CBT from psychodynamic therapy.[11] Common factors, such as therapeutic alliance, are thought to play a role in all psychotherapeutic treatments[29] and show a moderate association with outcomes.[30] Here, we found that "therapeutic praise" was positively associated with improvement, whereas "therapeutic empathy" showed a negative association. Rather than playing a causal role in outcomes, we believe increased empathy is likely to be indicative of a patient reporting a greater number of problems. Similarly, increased praise may be reflective of a patient responding well to treatment. Further research is required to determine the causal association between therapeutic alliance and outcomes, although previous work indicates therapeutic alliance may be reflective of a change in symptoms.[31]

We also investigated the association between patient variables and outcomes. Patient age (older patients showing better outcomes), absence of a long-term medical condition, not being prescribed psychotropic medication, and severity of anxiety symptoms were all positively associated with reliable improvement. By contrast, severity of depressive symptoms, the presence of a long-term medical condition, and being prescribed psychotropic medication were negatively associated. These results accord with previous work investigating treatment outcomes in a sample of approximately 3000 patients receiving IECBT.[32] Both studies report a positive association between GAD-7 scores and reliable improvement. Further work is needed to determine whether this reflects a greater association of CBT with short-term symptoms of anxiety and/or whether this effect may be specific to IECBT.

Limitations

A limitation of our approach is that it is not possible to determine whether a therapeutic feature is applied in an appropriate manner or whether a therapist adheres to the CBT protocol. It should be noted that the model provides a measure of the association between features and outcomes across sessions rather than measuring the quality of an individual session. Thus, future work needs to build on this approach to generate a validated model of session quality/adherence, alongside further refinement of the annotation guidelines and pooling of annotations. In addition, the model does not assess how the treatment was received by patients. To partly address this, we

are currently developing procedures to quantify patient utterances, enabling us to determine, for example, how use of change methods are associated with a change in patient's cognitions and whether therapeutic empathy is positively associated with outcomes after adjusting for the number of problems expressed by the patient.

We emphasize that our results only reveal the presence of an association between therapy content and outcomes, although some aspects of therapy (eg, change methods) are typically initiated by the therapist and appear likely to play a causal role. Further work is needed to determine the casual relationship between therapy features and outcomes by focusing on the temporal association between content and symptom change. Given the limited outcomes measures available, we are also unable to address the association between therapy content and long-term improvements in symptoms. In addition, other patient factors not included are likely to play a role in determining outcomes. Finally, it should be noted that for large data sets, the ORs and confidence intervals should be considered more informative of the clinical importance of a feature than statistical significance alone.

Conclusions

At present, the detailed monitoring of therapist performance requires expensive and time-consuming procedures. We believe that this work represents a first step toward a practicable approach for quality controlled behavioral health care. Such monitoring could help arrest therapist drift, ie, the failure to deliver treatments a therapist has been trained to deliver, which may be one of the biggest factors contributing to poor delivery of treatment.[33] Monitoring may help reverse the lower improvement rates observed in more experienced therapists.[34] We note that while a typical IAPT therapist may accrue substantial experience throughout a career (approximately 30 000 therapy hours), this data set represents an accumulation of knowledge from more than 90 000 hours of CBT. Deep leaning allows us to extract this knowledge to provide valuable insights into therapy that were previously unavailable to an individual therapist. As such, we believe this approach represents an important step in developing a data-driven understanding of mental health treatment and in improving the efficacy of psychotherapy.

ARTICLE INFORMATION

Accepted for Publication: July 18, 2019.

Published Online: August 22, 2019. doi:10.1001/jamapsychiatry.2019.2664

Author Contributions: Dr Ewbank had full access to all of the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis. Mr Martin created the utterance annotation software. Dr Cummins led the deep learning modeling.
Concept and design: All authors.
Acquisition, analysis, or interpretation of data: Ewbank, Cummins, Catarino, Martin, Blackwell.
Drafting of the manuscript: Ewbank.
Critical revision of the manuscript for important intellectual content: Cummins, Tablan, Bateup, Catarino, Martin, Blackwell.
Statistical analysis: Ewbank, Cummins.
Obtained funding: Tablan, Blackwell.
Administrative, technical, or material support: Cummins, Tablan, Bateup, Catarino, Martin, Blackwell.
Supervision: Tablan, Catarino, Blackwell.

Conflict of Interest Disclosures: All authors are employees of Ieso Digital Health. All authors report a patent to Methods and Systems for Improved Therapy Delivery and Monitoring pending.

Funding/Support: This study was funded by Ieso Digital Health.

Role of the Funder/Sponsor: As employees of Ieso Digital Health, the authors were responsible for the design and conduct of the study; collection, management, analysis, and interpretation of the data; preparation, review, or approval of the manuscript; and decision to submit the manuscript for publication.

Meeting Presentation: This paper was presented at the World Congress of Psychiatry; August 22, 2019; Lisbon, Portugal.

REFERENCES

1. Kilbourne AM, Beck K, Spaeth-Rublee B, et al. Measuring and improving the quality of mental health care: a global perspective. *World Psychiatry*. 2018;17(1):30-38. doi:10.1002/wps.20482

2. Johnsen TJ, Friborg O. The effects of cognitive behavioral therapy as an anti-depressive treatment is falling: a meta-analysis. *Psychol Bull*. 2015;141(4):747-768. doi:10.1037/bul0000015

3. Holmes EA, Ghaderi A, Harmer CJ, et al. The Lancet Psychiatry Commission on psychological treatments research in tomorrow's science. *Lancet Psychiatry*. 2018;5(3):237-286. doi:10.1016/S2215-0366(17)30513-8

4. Waller G, Turner H. Therapist drift redux: why well-meaning clinicians fail to deliver evidence-based therapy, and how to get back on track. *Behav Res Ther*. 2016;77:129-137. doi:10.1016/j.brat.2015.12.005

5. Perepletchikova F. On the topic of treatment integrity. *Clin Psychol (New York)*. 2011;18(2):148-153. doi:10.1111/j.1468-2850.2011.01246.x

6. Essock SM, Covell NH, Weissman EM. Inside the black box: the importance of monitoring treatment implementation. *Schizophr Bull*. 2004;30(3):613-615. doi:10.1093/oxfordjournals.schbul.a007107

7. Beck AT. Cognitive therapy: nature and relation to behavior therapy. *Behav Ther*. 1970;1:184-200. doi:10.1016/S0005-7894(70)80030-2

8. Ellis A. *Reason and Emotion in Psychotherapy*. New York: Lyle Stuart; 1962.

9. Perepletchikova F, Treat TA, Kazdin AE. Treatment integrity in psychotherapy research: analysis of the studies and examination of the associated factors. *J Consult Clin Psychol*. 2007;75(6):829-841. doi:10.1037/0022-006X.75.6.829

10. Barth J, Munder T, Gerger H, et al. Comparative efficacy of seven psychotherapeutic interventions for patients with depression: a network meta-analysis. *PLoS Med*. 2013;10(5):e1001454. doi:10.1371/journal.pmed.1001454

11. Blagys MD, Hilsenroth MJ. Distinctive activities of cognitive-behavioral therapy: a review of the comparative psychotherapy process literature. *Clin Psychol Rev*. 2002;22(5):671-706. doi:10.1016/S0272-7358(01)00117-9

12. Mulder R, Murray G, Rucklidge J. Common versus specific factors in psychotherapy: opening the black box. *Lancet Psychiatry*. 2017;4(12):953-962. doi:10.1016/S2215-0366(17)30100-1

13. Doss BD, Atkins DC. Investigating treatment mediators when simple random assignment to a control group is not possible. *Clin Psychol Sci Pract*. 2006;13:321-336. doi:10.1111/j.1468-2850.2006.00045.x

14. Bell EC, Marcus DK, Goodlad JK. Are the parts as good as the whole? a meta-analysis of component treatment studies. *J Consult Clin Psychol*. 2013;81(4):722-736. doi:10.1037/a0033004

15. Lorenzo-Luaces L, German RE, DeRubeis RJ. It's complicated: the relation between cognitive change procedures, cognitive change, and symptom change in cognitive therapy for depression. *Clin Psychol Rev*. 2015;41:3-15. doi:10.1016/j.cpr.2014.12.003

16. Kessler D, Lewis G, Kaur S, et al. Therapist-delivered internet psychotherapy for depression in primary care: a randomised controlled trial. *Lancet*. 2009;374(9690):628-634. doi:10.1016/S0140-6736(09)61257-5

17. National Institute for Health and Clinical Excellence. Common mental health disorders: the NICE guideline on identification and pathways to care. London, England: RCPsych Publication; 2011.

18. Roth AD, Pilling S. Using an evidence-based methodology to identify the competences required to deliver effective cognitive and behavioural therapy for depression and anxiety disorders.. *Behav Cogn Psychother*. 2008;36(February):129-147. doi:10.1017/S1352465808004141

19. Clark DM. Implementing NICE guidelines for the psychological treatment of depression and anxiety disorders: the IAPT experience. *Int Rev Psychiatry*.

2011;23(4):318-327. doi:10.3109/09540261.2011.606803

20. N.I.C.E. Principles for Best Practice in Clinical Audit. Radcliffe Medical Press; 2002. http://www.uhbristol.nhs.uk/files/nhs-ubht/best_practice_clinical_audit.pdf. Accessed January 31, 2019.

21. Kroenke K, Spitzer RL, Williams JB. The PHQ-9: validity of a brief depression severity measure. *J Gen Intern Med*. 2001;16(9):606-613. doi:10.1046/j.1525-1497.2001.016009606.x

22. Spitzer RL, Kroenke K, Williams JB, Löwe B. A brief measure for assessing generalized anxiety disorder: the GAD-7. *Arch Intern Med*. 2006;166(10):1092-1097. doi:10.1001/archinte.166.10.1092

23. Blackburn I, James IA, Milne DL, et al. The revised cognitive therapy scale (CTS-R): psychometric properties. *Behav Cogn Psychother*. 2001;29:431-446. doi:10.1017/S1352465801004040

24. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics*. 1977;33(1):159-174. doi:10.2307/2529310

25. Alford BA, Beck AT. Cognitive therapy of delusional beliefs. *Behav Res Ther*. 1994;32(3):369-380. doi:10.1016/0005-7967(94)90134-1

26. Kazantzis N, Whittington C, Zelencich L, Kyrios M, Norton PJ, Hofmann SG. Quantity and quality of homework compliance: a meta-analysis of relations with outcome in cognitive behavior therapy. *Behav Ther*. 2016;47(5):755-772. doi:10.1016/j.beth.2016.05.002

27. Beck JS. *Cognitive Behavior Therapy: Basics and Beyond*. 2nd ed. New York, NY: Guilford Press; 2011.

28. Alford BA, Beck AT. *The Integrative Power of Cognitive Therapy*. New York, NY: Guilford Press; 1997. doi:10.1891/0889-8391.11.4.309

29. Wampold BE. How important are the common factors in psychotherapy? an update. *World Psychiatry*. 2015;14(3):270-277. doi:10.1002/wps.20238

30. Martin DJ, Garske JP, Davis MK. Relation of the therapeutic alliance with outcome and other variables: a meta-analytic review. *J Consult Clin Psychol*. 2000;68(3):438-450. doi:10.1037/0022-006X.68.3.438

31. Strunk DR, Brotman MA, DeRubeis RJ. The process of change in cognitive therapy for depression: predictors of early inter-session symptom gains. *Behav Res Ther*. 2010;48(7):599-606. doi:10.1016/j.brat.2010.03.011

32. Catarino A, Bateup S, Tablan V, et al. Demographic and clinical predictors of response to internet-enabled cognitive-behavioural therapy for depression and anxiety. *BJPsych Open*. 2018;4(5):411-418. doi:10.1192/bjo.2018.57

33. Waller G. Evidence-based treatment and therapist drift. *Behav Res Ther*. 2009;47(2):119-127. doi:10.1016/j.brat.2008.10.018

34. Shapiro DA, Shapiro D. Meta-analysis of comparative therapy outcome studies: a replication and refinement. *Psychol Bull*. 1982;92(3):581-604. doi:10.1037/0033-2909.92.3.581

Appendix 2: Catarino, Bateup, Tablan, Innes, Freer, Richards, Stott, Hollon, Chamberlain, Hayes and Blackwell (2018)

BJPsych open

BJPsych Open (2018)
4, 411–418. doi: 10.1192/bjo.2018.57

Demographic and clinical predictors of response to internet-enabled cognitive–behavioural therapy for depression and anxiety

Ana Catarino, Sarah Bateup, Valentin Tablan, Katherine Innes, Stephen Freer, Andy Richards, Richard Stott, Steven D. Hollon, Samuel Robin Chamberlain, Ann Hayes and Andrew D. Blackwell

Background

Common mental health problems affect a quarter of the population. Online cognitive–behavioural therapy (CBT) is increasingly used, but the factors modulating response to this treatment modality remain unclear.

Aims

This study aims to explore the demographic and clinical predictors of response to one-to-one CBT delivered via the internet.

Method

Real-world clinical outcomes data were collected from 2211 NHS England patients completing a course of CBT delivered by a trained clinician via the internet. Logistic regression analyses were performed using patient and service variables to identify significant predictors of response to treatment.

Results

Multiple patient variables were significantly associated with positive response to treatment including older age, absence of long-term physical comorbidities and lower symptom severity at start of treatment. Service variables associated with positive response to treatment included shorter waiting times for initial assessment and longer treatment durations in terms of the number of sessions.

Conclusions

Knowledge of which patient and service variables are associated with good clinical outcomes can be used to develop personalised treatment programmes, as part of a quality improvement cycle aiming to drive up standards in mental healthcare. This study exemplifies translational research put into practice and deployed at scale in the National Health Service, demonstrating the value of technology-enabled treatment delivery not only in facilitating access to care, but in enabling accelerated data capture for clinical research purposes.

Declaration of interest

A.C., S.B., V.T., K.I., S.F., A.R., A.H. and A.D.B. are employees or board members of the sponsor. S.R.C. consults for Cambridge Cognition and Shire. Keywords: Anxiety disorders; cognitive behavioural therapies; depressive disorders; individual psychotherapy

Common mental health disorders including depression and anxiety are characterised by intense emotional distress and have an impact on social and occupational functioning. One in four adults are estimated to have a mental health problem in any given year[1,2] but a significant proportion remain undiagnosed or untreated.[3] The Improving Access to Psychological Therapies (IAPT) programme is a large-scale initiative aimed at increasing access to accountable, evidence-based psychological therapy for common mental health disorders within the English National Health Service (NHS), while controlling costs.[4] In 2015/2016 one-third of patients referred to IAPT received cognitive–behavioural therapy (CBT).[5] With proven clinical effectiveness, structured CBT models are also amenable to delivery via 'self-help' programmes including online materials,[6,7] with various online CBT models shown as efficacious in controlled trial settings.[6–8] Online CBT offers a flexibility that is not possible under face-to-face programmes, allowing patients to undergo treatment outside of office hours, optimising convenience and avoiding perceived stigma associated with undertaking in-person therapy.[9] Other potential benefits include shorter waiting times, enhanced disclosure and improved access for patients who are reluctant to contact services given the nature of their condition (for example agoraphobia), or cannot travel because of disability or geographical location.[9,10] Despite multiple potential advantages, varying degrees of therapist intervention in online CBT can affect clinical outcomes and patient engagement in real-world settings,[5,8,11,12] with therapist-guided online CBT associated with better clinical outcomes and lower drop-out rates than self-guided programmes. Internet-enabled CBT (IECBT) is an effective type of online therapy,[13] successfully used within IAPT, whereby patients are offered weekly one-to-one sessions with an accredited CBT therapist. As a result of its one-to-one nature, IECBT is classed as a high-intensity therapy and can be used to treat more severe patients, relative to other self-guided and guided self-help online CBT modalities that are classed as low-intensity interventions and therefore only suitable for patients with milder presentations.[14] Previous research investigating predictors of clinical outcomes for low-intensity guided self-help interventions shows that higher levels of adherence to treatment and treatment credibility are associated with higher rates of improvement and lower post-treatment scores.[15,16] This highlights the importance of investigating predictors of clinical outcomes in response to high-intensity online interventions like IECBT, where the synchronous, yet anonymous, nature of the interaction between therapist and patient may promote treatment credibility and patient adherence to treatment protocol. The present study aims to explore patient and service variables as predictors of clinical outcomes, in patients receiving IECBT for the treatment of a common mental health disorder.

Method

Data were analysed from patients receiving IECBT for the treatment of a mental health disorder, between April 2015 and March 2016. IECBT was delivered using a commercial package, originally developed for and currently used in the English NHS, provided by Ieso Digital Health (http://uk.iesohealth.com). Patients self-referred or were referred by a primary healthcare worker directly to the service in the regions of Surrey, West Kent, Camden and East Riding of Yorkshire. Patients can register with the service using an online registration form or over the phone. Patients reporting suicidal intent, during registration or at any point during the episode of care, were appropriately advised online by their therapist or another member of the clinical team, and signposted to specialist services accordingly. In exceptional circumstances of immediate or serious risk patients were contacted over the phone by their therapist or a clinical supervisor.

After registration, patients were assigned to a qualified CBT therapist accredited by the British Association for Behavioural and Cognitive Psychotherapies. Initial assessments were carried out in an online therapy room via one-to-one real-time written conversation, after which the therapist assigned the patient a diagnosis and National Institute for Health and Care Excellence-approved disorder-specific CBT treatment protocols,[14] based on Roth & Pilling's CBT competences framework,[17] were delivered during weekly sessions. Treatment duration was determined by the therapist based on their clinical judgement, with typical treatment length between 6 and 8 sessions. Between treatment appointments, asynchronous messages and homework tasks were exchanged between therapist and patient, promoting engagement and adherence to evidence-based treatment models. All communication between therapist and patient was done exclusively online through Ieso's proprietary platform, following internationally recognised standards for information security (ISO 27001; https://www.iesohealth.com/en-gb/legal/iso-certificates).

Clinical outcomes in the IECBT group were referenced against reported outcomes for patients referred to IAPT services in the same time period and same regions where IECBT was offered. Patients in the reference group received care as usual, comprising high- and low-intensity treatments, face-to-face and online therapy services, including IECBT.[18]

The information captured through IAPT's minimum data-set, including IECBT, is intended to support monitoring of implementation and effectiveness of national policy/legislation, policy development, performance analysis and benchmarking, national analysis and statistics and national audit of IAPT services. At registration patients agree to the services' terms and conditions, including use of anonymised data for audit purposes and to support research, including academic publications or conference presentations.

Outcomes measures

Clinical outcomes were measured in terms of clinical recovery and improvement, defined following IAPT guidelines.[4,19] According to IAPT convention, these measures are defined for patients undergoing a minimum of two sessions of therapy. This is the minimum dose of therapy a patient must receive such that pre- and post-treatment scores are collected and clinical change can be estimated.[19] Clinical recovery and improvement are calculated based on two severity measures completed by the patient at initial assessment and for every therapy session (completion rate 95%): Patient Health Questionnaire (PHQ)-9[20] and General Anxiety Disorder (GAD)-7,[21] corresponding to depressive and anxiety symptoms, respectively.

The PHQ-9 is a nine-item measure designed to facilitate screening and severity assessment of depression, ranging from 0 to 27 and with a recommended cut-off of ten or more for distinguishing patients considered to have clinically significant depressive symptoms.[20] A reduction of six points or more on the PHQ-9 scale between two time points is indicative of statistically reliable improvement in symptom severity.[20]

The GAD-7 is a seven-item screening and severity measure for generalised anxiety disorder, ranging from 0 to 21 and with a recommended cut-off of eight or more for distinguishing patients considered to have clinically significant anxiety symptoms.[21,22] A reduction of four points or more on the GAD-7 scale between two time points is indicative of statistically reliable improvement in symptom severity.[21]

If a patient scores above the clinical threshold for one or both of these measures at initial assessment (i.e. ten or above for PHQ-9 and/or eight or above for GAD-7), they are classed as meeting 'caseness' at assessment. Other symptom severity measures, such as severity scores for subtypes of anxiety disorders, were not examined as only PHQ-9 and GAD-7 are mandatorily collected within the IAPT framework.

For patients undergoing two or more therapy sessions, the difference between scores at initial assessment and last treatment session for PHQ-9 and GAD-7 is used to determine patients' recovery status. Recovery is a binary measure. Under IAPT guidelines, patients with two or more therapy sessions who move from above caseness at assessment to below caseness on both the PHQ-9 and GAD-7 scales at the last treatment session are classed as recovered. The recovery rate for a group of patients is calculated as number of patients recovered, divided by number of patients at caseness at initial assessment.

Improvement is also a binary measure. Under IAPT guidance patients with two or more therapy sessions who show a significant reduction in at least one of the outcome measures from assessment to the last treatment session, while not showing a significant increase in the other outcome measure, were classed as improved (i.e. decrease of six points or more in the PHQ-9 and/or four points or more in the GAD-7, while not simultaneously showing an increase of six points or more in the PHQ-9 or four points or more in the GAD-7). The improvement rate for a group of patients is calculated as number of patients improved, divided by number of patients with two or more therapy sessions. Patients who simultaneously improve and recover are classed as reliably recovered.

Sample size

A total of 4468 patients registered with the IECBT service between April 2015 and March 2016. Of these, 487 patients (11%) did not meet the eligibility criteria (over 18 years old, registered with a general practitioner in the geographical region where the service is commissioned) and were signposted to other mental health services as appropriate. From the remaining 3981 eligible patients, 95 (2%) were deemed not suitable for the service for clinical reasons (for example risk, Axis II disorder) and were signposted to other mental health services as appropriate. A total of 3886 patients were offered treatment, of which 2211 (57%) had two or more treatment sessions. Out of these 2211 patients, 1818 (82%) were at caseness at assessment (170 at caseness according to PHQ-9 only, 375 at caseness for GAD-7 only and 1273 at caseness for both) (Fig. 1). A comparison of demographics between patients undergoing two or more therapy sessions and patients who drop-out before this point can be found in the supplementary materials (Supplementary Table 1, available at https://doi.org/10.1192/bjo.2018.57).

Between April 2015 and March 2016 a total of 45 560 referrals were received by IAPT services in the same regions where IECBT was offered. In total, 19 325 patients were discharged in this time

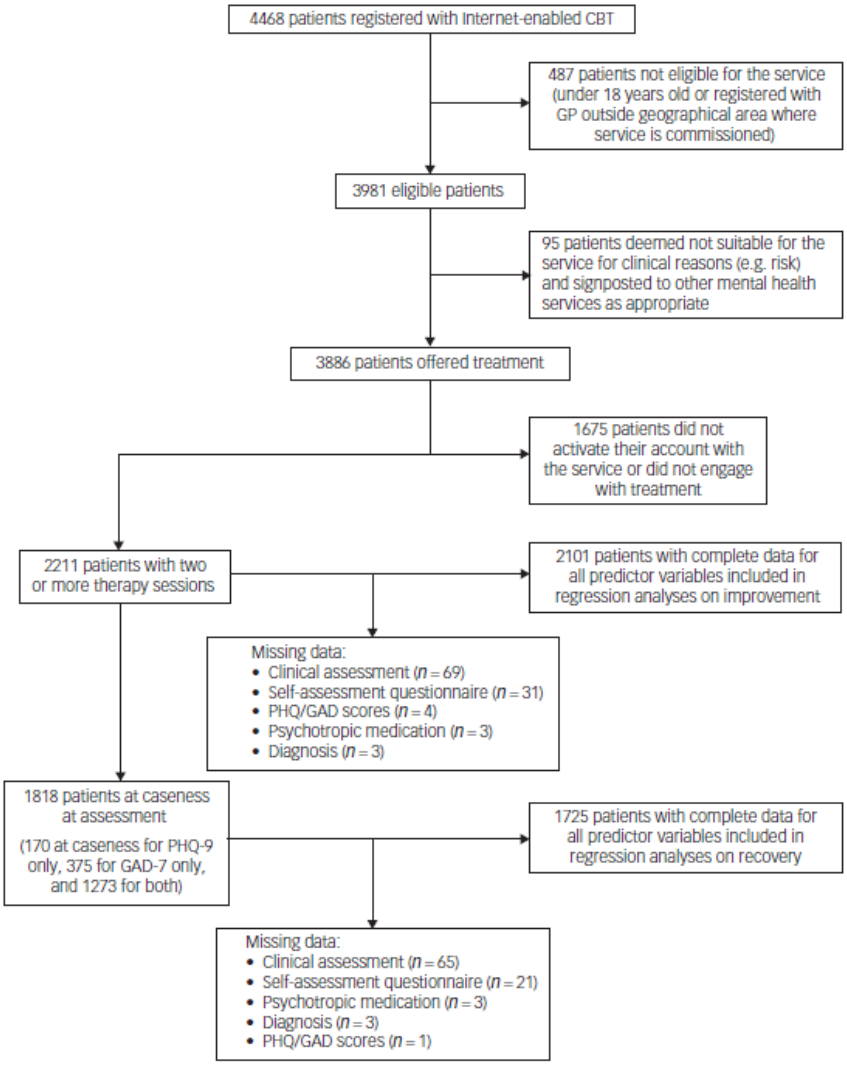


Fig. 1 Study profile and patient flow chart.

GP, general practitioner; PHQ, Patient Health Questionnaire; GAD General Anxiety Disorder.

period having had two therapy sessions or more, of which 17 470 (90%) were at caseness at assessment. Data analyses focusing on the improvement metric were conducted on data from patients with two or more therapy sessions, while analyses focusing on the recovery metric were conducted on data from patients at caseness at assessment, who also had two or more therapy sessions.

Regression analyses – predictors of clinical response in IECBT

Logistic regression analyses were performed in R to identify significant predictors of recovery and improvement in patients receiving IECBT, based on patient demographics and service variables. Demographic variables included patient gender, age, severity, diagnosis, whether or not the patient had a long-term physical condition, and whether or not the patient was taking psychotropic medication (for example antidepressants or anxiolytics) at the start of treatment. Service variables comprised data pertaining to a patient's course of treatment, including waiting times between various stages in the patient journey, treatment duration and number of scheduled appointments the patient failed to attend.

Given the nature of the statistical models employed, record sets were included only for cases with complete data for all predictor variables. Of the initial sample of 2211 patients with two therapy sessions and 1818 patients at caseness at assessment, 95% had

complete data for all predictor variables and were included in the analyses ($n = 2101$ for improvement analysis, $n = 1725$ for recovery analysis; Fig. 1).

Continuous predictor variables were scaled and centred to the mean. Multicollinearity analyses were performed to investigate potential correlations between predictor variables. Statistical significance was defined as $P<0.05$ two-tailed, uncorrected.

Comparative analysis of clinical outcomes

Although inferential analysis of comparative clinical effectiveness is not possible in the present study because of the lack of a face-to-face control group, publicly available IAPT data makes it possible to reference IECBT clinical outcomes against averages for the same time period and geographical regions. Patients with two or more therapy sessions in the IECBT group ($n = 2211$) were matched to the IAPT reference group ($n = 19\,325$) for severity (PHQ-9 and GAD-7 scores at assessment), using a multivariate rejection sampling algorithm implemented in R.[23,24] Lack of publicly available distribution data for other variables means it was not possible to match the two groups for other potentially relevant variables such as age, diagnosis and presence of long-term physical comorbidities. Clinical outcomes of the severity-matched IECBT group relative to IAPT are reported.

Results

Regression analyses – predictors of clinical response in IECBT

Logistic regression analyses identified presence of long-term physical conditions, initial GAD-7 scores, waiting time for assessment, total number of treatment sessions and patient age as significant predictors of improvement (Table 1). Apart from waiting time for assessment, these variables were also identified as significant predictors of recovery, in addition to initial PHQ-9 scores (Table 2).

Results show that patients with long-term physical conditions are less likely to show good clinical outcomes compared with patients without long-term conditions (Tables 1 and 2). Patients with higher severity scores at assessment are also less likely to show clinical recovery (Table 2). However, in line with what can be observed from Table 1, results suggest that patients with higher GAD-7 scores at assessment have a higher likelihood of showing clinical improvement.

A significant positive association between patient age and likelihood of good clinical outcomes was also observed. This association was explored further in a *post hoc* analysis that revealed a significant negative correlation between patient age and severity (age and PHQ-9: $r = -0.09$, $t = -4.02$, d.f. = 2102, $P<0.001$; age and GAD-7: $r = -0.13$, $t = -5.98$, d.f. = 2102, $P<0.001$), as well as a weak but significant positive correlation between patient age and number of treatment sessions ($r = 0.05$, $t = 2.10$, d.f. = 2102, $P=0.036$).

Finally, results show that patients who have undergone a larger number of therapy sessions are more likely to show good clinical outcomes (Tables 1 and 2). However, *post hoc* analyses showed no significant association between treatment duration and clinical outcomes in patients with five or more sessions. Clinical outcome rates were optimal and less variable for treatment durations of five to nine sessions (51% of patients with more than two sessions, recovery rate: 57–60%; improvement rate: 67–72%). Clinical outcomes for patients with more than two but fewer than five treatment sessions are significantly lower and more variable (14% of all patients with

Table 1 Results of logistic regression analysis investigating predictors of improvement in the Internet-enabled cognitive–behavioural therapy cohort ($n = 2101$)[a]

Predictor variable	Mean/prevalence	*b*	s.e.	Wald's statistic, z^2	*P*	Subgroup *n*	Improvement rate, %
Gender, %							
Men	27.3	–	–	–	–	574	60.1
Women	72.1	0.04	0.11	0.15	0.698	1514	62.2
Unknown/not stated	0.6	−0.24	0.59	0.17	0.682	13	53.8
Age, years: mean	36.8	0.12	0.05	5.87	0.015*	–	–
Long-term physical conditions, %							
No	34.3	–	–	–	–	721	60.9
Yes	20.8	−0.32	0.14	5.29	0.021*	436	59.9
Unknown/not stated	44.9	0.02	0.11	0.05	0.819	944	62.9
Diagnosis,[b] %							
Anxiety	42.1	–	–	–	–	885	61.4
Depression	22.6	0.03	0.13	0.04	0.851	474	60.1
Other	35.3	0.13	0.11	1.30	0.254	742	62.8
Psychotropic medication, %							
Prescribed not taking	8.1	–	–	–	–	171	59.6
Prescribed taking	38.3	0.24	0.18	1.78	0.182	805	63.6
Not prescribed	52.5	0.26	0.18	2.05	0.152	1103	60.4
Unknown/not stated	1.0	0.23	0.49	0.21	0.649	22	63.6
StartPhq9, mean	12.7	0.05	0.07	0.50	0.479	–	–
StartGad7, mean	11.9	0.53	0.06	70.91	<0.001***	–	–
WaitingSAQ, days: mean	3.0	0.01	0.06	0.02	0.894	–	–
WaitingAssignment, days: mean	0.8	−0.04	0.06	0.44	0.508	–	–
WaitingContact, days: mean	1.1	−0.07	0.08	0.93	0.336	–	–
WaitingAssessment, days: mean	7.6	−0.14	0.05	7.54	0.006**	–	–
WaitingTreatment, days: mean	8.9	−0.02	0.05	0.18	0.671	–	–
NumberSessions, sessions: mean	5.5	0.29	0.05	32.09	<0.001***	–	–
NumberDNA, sessions: mean	0.5	−0.01	0.05	0.06	0.813	–	–

StartPhq9, Patient Health Questionnaire-9 score at assessment; StartGad7, General Anxiety Disorder-7 score at assessment; WaitingSAQ, time between referral and patient completing self-assessment questionnaire; WaitingAssignment, waiting time from patient completing the self-assessment questionnaire and therapist assignment; WaitingContact, waiting time between therapist assignment and first contact by the therapist; WaitingAssessment, waiting time between first contact from the therapist and clinical assessment appointment; WaitingTreatment, waiting time between clinical assessment and first therapy session; NumberSessions, total number of therapy sessions attended by the patient; NumberDNA, number of scheduled appointments the patient failed to attend.
a. Men, long-term physical conditions 'no', diagnosis 'anxiety' and psychotropic medication 'prescribed not taking' were reference classes for the categorical variables.
b. The diagnosis anxiety encompasses patients diagnosed with agoraphobia, generalised anxiety disorder, hypochondriacal disorder, obsessive–compulsive disorder, panic disorder, post-traumatic stress disorder, social phobia, specific phobia or anxiety disorder unspecified. The diagnosis depression encompasses patients diagnosed with depressive episode, dysthymia or recurrent depressive disorder. The diagnosis 'other' encompasses all diagnoses not otherwise listed (for example chronic pain, bereavement, eating disorders).
*P<0.05, **P<0.01, ***P<0.001.

Table 2 Results of logistic regression analysis investigating predictors of recovery in the internet-enabled cognitive–behavioural therapy cohort ($n = 1725$)[a]

Predictor variable	Mean/prevalence	*b*	s.e.	Wald's statistic, z^2	*P*	Subgroup, *n*	Recovery rate, %
Gender, %							
Men	26.5	–	–	–	–	457	51.0
Women	73.0	0.15	0.12	1.62	0.202	1259	50.8
Unknown/not stated	0.5	−0.14	0.70	0.04	0.846	9	44.4
Age, years: mean	36.3	0.23	0.06	18.05	<0.001***	–	–
Long-term physical conditions, %							
No	33.2	–	–	–	–	573	55.8
Yes	22.0	−0.37	0.15	5.88	0.015*	379	43.5
Unknown/not stated	44.8	−0.14	0.12	1.32	0.251	773	50.6
Diagnosis, %[b]							
Anxiety	40.8	–	–	–	–	703	53.3
Depression	23.6	−0.04	0.15	0.07	0.796	407	45.9
Other	35.7	0.02	0.13	0.02	0.875	615	51.1
Psychotropic medication, %							
Prescribed not taking	8.8	–	–	–	–	151	47.0
Prescribed taking	41.7	0.01	0.20	0.01	0.942	719	46.3
Not prescribed	48.3	0.06	0.20	0.09	0.769	834	55.6
Unknown/not stated	1.2	−0.49	0.53	0.84	0.360	21	38.1
StartPhq9, mean	14.3	−0.56	0.07	70.32	<0.001***	–	–
StartGad7, mean	13.4	−0.30	0.06	24.76	<0.001***	–	–
WaitingSAQ, days: mean	3.0	0.01	0.06	0.02	0.900	–	–
WaitingAssignment, days: mean	0.8	0.04	0.06	0.36	0.550	–	–
WaitingContact, days: mean	1.2	−0.11	0.07	2.16	0.142	–	–
WaitingAssessment, days: mean	7.4	−0.08	0.06	2.16	0.141	–	–
WaitingTreatment, days: mean	8.8	−0.02	0.05	0.18	0.674	–	–
NumberSessions, sessions: mean	5.6	0.32	0.05	33.83	<0.001***	–	–
NumberDNA, sessions: mean	0.5	−0.11	0.05	3.81	0.051	–	–

StartPhq9, Patient Health Questionnaire-9 score at assessment; StartGad7, General Anxiety Disorder-7 score at assessment; WaitingSAQ, time between referral and patient completing self-assessment questionnaire; WaitingAssignment, waiting time from patient completing the self-assessment questionnaire and therapist assignment; WaitingContact, waiting time between therapist assignment and first contact by the therapist; WaitingAssessment, waiting time between first contact from the therapist and clinical assessment appointment; WaitingTreatment, waiting time between clinical assessment and first therapy session; NumberSessions, total number of therapy sessions attended by the patient; NumberDNA, number of scheduled appointments the patient failed to attend.
a. Men, long-term physical conditions 'no', diagnosis 'anxiety' and psychotropic medication 'prescribed not taking' were reference classes for the categorical variables.
b. The diagnosis anxiety encompasses patients diagnosed with agoraphobia, generalised anxiety disorder, hypochondriacal disorder, obsessive–compulsive disorder, panic disorder, post-traumatic stress disorder, social phobia, specific phobia or anxiety disorder unspecified. The diagnosis depression encompasses patients diagnosed with depressive episode, dysthymia or recurrent depressive disorder. The diagnosis 'other' encompasses all diagnoses not otherwise listed (for example chronic pain, bereavement, eating disorders).
*P<0.05, ***P<0.001.

more than two sessions, recovery rate: 27–53%; improvement rate: 42–61%).

Tests of the full models against constant-only models were significant for both regression analyses (improvement regression model: $\chi^2(19) = 195.95$, $P<0.001$; recovery regression model: $\chi^2(19) = 278.36$, $P<0.001$). Additionally, multicollinearity analyses revealed variance inflation factors smaller than two for all predictor variables. This is the standard threshold value for indicating the presence of multicollinearity in this type of analysis, thus confirming that regression models were not affected by the presence of multicollinearity.[25]

IAPT's improvement and recovery metrics are, by definition, biased by initial symptom severity, which confounds interpretation of the results. In a *post hoc* regression analysis investigating predictors of clinical outcomes while controlling for artefactual relations with initial severity, we defined per cent improvement as a 25% decrease in scores for both the PHQ-9 and the GAD-7.[26] Similar to IAPT convention, we classed a patient as achieving per cent improvement if they showed a 25% decrease in scores in one or both scales, without showing deterioration in either scale. Results of this analysis show broadly similar results to the analysis on predictors of improvement as defined according to IAPT convention, but they no longer show the significant association between initial severity scores and per cent improvement (Supplementary Table 2).

Following IAPT convention, improvement, recovery and per cent improvement metrics were defined combining the PHQ-9 and GAD-7 scales. Although this allows for a more comprehensive characterisation of the patients, who often present with a combination of depressive and anxiety features, these are two separate scales, measuring different constructs. It can be hypothesised that patient and service variables may have an impact on the likelihood of good clinical outcomes for each scale differently. *Post hoc* regression analyses investigating predictors of per cent improvement for each scale separately are presented in Supplementary Tables 3 and 4.

Clinical outcomes adjusted for symptom severity and benchmarked against national audit comparator data

Variations were observed in the likelihood of improvement and recovery with PHQ-9 and GAD-7 scores at assessment, for patients treated with IECBT. Regression analyses results show that patients with higher PHQ-9 and GAD-7 scores at assessment, have lower likelihoods of recovery but equivalent or higher likelihood for improvement. Accordingly, IECBT clinical improvement and recovery data were benchmarked against severity-matched cohorts. Severity-matched patients treated with IECBT showed similar improvement and recovery rates relative to IAPT patients, as well as similar magnitude of symptom reduction, pre- and post-treatment (Table 3). Although classical significance testing was avoided because of bias in favour of rejecting the null hypothesis for large sample sizes,[27] effect sizes and 95% confidence intervals are presented (Table 3). Despite some isolated differences in disorder distribution across the two cohorts, the observed odds ratios and effect sizes are generally small for most variables.[28] Together with differences in clinical outcomes of less than 1% and differences in magnitude of symptom reduction of 0.6 points or less between the two groups, these results suggest that differences between the

Table 3 Demographic details and clinical outcomes for patients finishing a course of treatment between April 2015 and March 2016[a]

	IECBT (severity matched) (n = 2211)		IAPT (areas where IECBT was offered) (n = 19 325)		Group difference	
	Mean (s.d.) 95% CI	% (95% CI)	Mean (s.d.) 95% CI	%, 95% CI	Cohen's d (95% CI)	OR (95% CI)
Gender, women	–	73 (71–75)	–	67 (67–68)	–	1.30 (1.18 to 1.44)
Long-term physical conditions	–	21 (20–23)	–	23 (22–23)	–	0.92 (0.83 to 1.03)
Diagnosis[b]						
Depression	–	27 (25–29)	–	36 (35–36)	–	0.68 (0.62 to 0.75)
Anxiety	–	41 (39–43)	–	38 (37–39)	–	1.13 (1.03 to 1.24)
Other	–	32 (30–34)	–	27 (26–27)	–	1.30 (1.18 to 1.43)
PHQ-9 at assessment	13.8 (4.5) 13.6 to 14.0	–	14.3 (5.2) 14.2 to 14.4	–	−0.10 (−0.14 to −0.05)	–
GAD-7 at assessment	13.0 (3.8) 12.8 to 13.1	–	13.2 (4.2) 13.1 to 13.3	–	−0.05 (−0.09 to 0.004)	–
PHQ-9 post-treatment	8.6 (6.1) 8.3 to 8.9	–	8.5 (6.2) 8.4 to 8.6	–	0.02 (−0.03 to 0.06)	–
GAD-7 post-treatment	8.0 (5.4) 7.8 to 8.2	–	7.8 (5.4) 7.7 to 7.9	–	0.04 (−0.01 to 0.08)	–
Treatment length, sessions	5.5 (2.8) 5.4 to 5.6	–	6.4[c]	–	N/A	–
Improvement rate	–	65.8 (63.8 to 67.8)	–	66.2 (65.5 to 66.8)	–	0.98 (0.90 to 1.08)
Recovery rate	–	48.8 (46.7 to 51.0)	–	49.5 (48.8 to 50.3)	–	0.97 (0.89 to 1.06)

CBT, cognitive–behavioural therapy; IECBT, internet-enabled CBT; IAPT, Improving Access to Psychological Therapies; PHQ, Patient Health Questionnaire; GAD, General Anxiety Disorder; NA, not available.
a. The IECBT cohort was severity matched to the IAPT cohort using a multivariate rejection sampling.
b. The diagnosis anxiety encompasses patients diagnosed with agoraphobia, generalised anxiety disorder, hypochondriacal disorder, obsessive–compulsive disorder, panic disorder, post-traumatic stress disorder, social phobia, specific phobia or anxiety disorder unspecified. The diagnosis depression encompasses patients diagnosed with depressive episode, dysthymia or recurrent depressive disorder. The diagnosis 'other' encompasses all diagnoses not otherwise listed (for example chronic pain, bereavement, eating disorders).
c. Average treatment duration for all conditions nationwide, regional data not available.

two groups in clinical outcomes and score reduction are unlikely to be meaningful.[29]

Discussion

This paper reports the first real-world (non-randomised controlled trial) clinical outcomes data for patients with depression and anxiety treated using internet-enabled CBT. The first evidence of clinical efficacy of IECBT in depression was published in the Lancet in 2009.[13] These extended data, including both depression and anxiety disorders, offer an example of translational research put into practice and successfully deployed at scale. The application of the resultant data-set in advancing our understanding of clinical and demographic variables associated with response to treatment suggests that there might be value in data-enabled mental health services as platforms for clinical research. Knowledge acquired with these tools can be used to refine service specifications and develop personalised treatment programmes, as part of a quality improvement cycle aiming to drive up standards in mental healthcare.

Main findings

Regression analyses revealed a significant association between initial psychometric scores and likelihood of recovery, with higher scores associated with lower recovery rates. By definition, patients recover by going below the caseness threshold for both PHQ-9 and GAD-7. Therefore, it is not unexpected that patients whose initial scores are closer to that threshold have higher chances of recovery. This does, however, raise the question of whether recovery alone is a suitable index to measure clinical outcome.

As indicated by Gyani and colleagues[19] the recovery metric does not take into account whether the observed reduction in severity is greater than the measurement error of the scales. Conversely the improvement index is a measure of whether or not a reduction in severity is statistically reliable, regardless of caseness and may therefore be a better metric for widespread use. In the present study, patients with higher initial scores are more likely to show clinical improvement, as validated by the results of the regression analysis, where patients with higher initial GAD-7 scores show higher likelihoods of improvement (Table 1). IAPT's reliable recovery index is a composite metric measuring whether a patient recovered while simultaneously showing a statistically reliable reduction in severity. Although by definition this metric may be less susceptible to bias in favour of patients who are near the recovery threshold, it will still be biased against patients with higher severity scores at assessment, who will be less likely to cross the recovery threshold.

To investigate predictors of clinical outcomes while controlling for artefactual relations with initial severity, we conducted a *post hoc* analysis investigating predictors of per cent improvement.[26] Results show broadly similar results to the regression analysis on predictors of improvement as defined according to IAPT convention, but the significant association with initial severity scores is no longer present with the per cent improvement measure (Supplementary Table 2). Although differences in recovery rate with severity may be expected in this context, they may also indicate the presence of non-specific treatment effects. Future strategies to improve treatment effectiveness should therefore be aimed at boosting recovery of more severe patients, including increased session frequency at the start of treatment, or the use of specific CBT protocols for severe depression.

Regression analyses on improvement and recovery also revealed significant associations between clinical outcomes and age, presence of a long-term physical condition and number of therapy sessions. Results show that greater age is associated with better clinical

outcomes, in contrast with previous research showing lower effectiveness of CBT in older adults.[30] However, it is important to note that in the present study the mean age of the patient cohort was 36 years, whereas previous research on the effects of CBT on older adults focused on adults over the age of 55.[30] Older adults are more likely to be affected by age-related cognitive decline and physical comorbidities that may directly influence CBT outcomes but are not prevalent factors in the current cohort.

Post hoc analyses on predictors of per cent improvement reveal age to be a positive predictor of likelihood of per cent improvement, similar to what was observed for the analysis on improvement defined under IAPT's convention. This suggests that despite a significant negative correlation between patient age and severity, the association between age and clinical outcomes is not in this case driven by differences in severity across the age range in this particular cohort. A weak but significant positive correlation between patient age and number of treatment sessions, as well as higher mean age of patients with two or more therapy sessions (Supplementary Table 1) suggests that in this particular cohort, older patients may be less likely to drop-out at earliest stages of treatment, therefore benefitting from a larger therapeutic dose and consequently be more likely to achieve positive clinical outcomes.

In this cohort it was also observed that patients with long-term physical conditions were less likely to show good clinical outcomes. This finding is unsurprising given that long-term physical conditions are often associated with comorbid mental health problems and complex psychological issues, which may themselves be chronic in nature and often treatment resistant.[31,32] Lower probability of response to treatment may signal the need for tailored, condition-specific CBT models, so patients can be helped to deal with mental and physical symptoms in an integrated fashion. However, it could also be argued that PHQ-9 and GAD-7, used to calculate clinical outcomes, lack sensitivity to detect clinically significant improvements in patients with long-term conditions. Disease-specific measures that may better reflect the complexities of the physical disease and provide better indicators of clinical improvement in these patients, should be included in future studies.

Service variables shown to be associated with the likelihood of good clinical outcomes included higher number of therapy sessions and reduced waiting time for assessment. Although these findings are supported by similar reports in the literature,[33] care should be taken when drawing causal inferences. At first glance these results convey the impression that longer courses of treatment are associated with better clinical outcomes. However, an alternative explanation is that patients who do not adhere to their treatment plan and drop-out at an earlier stage during the course of treatment, therefore receiving a subtherapeutic dose, are less likely to achieve good clinical outcomes. This hypothesis is supported by results of *post hoc* analyses showing that clinical outcome rates were optimal and less variable for treatment durations of five to nine sessions, whereas patients with more than two but fewer than five treatment sessions observed significantly lower and more variable clinical outcomes. Difficulties with engagement leading to poor clinical outcomes may be particularly relevant in patients with more severe depressive symptoms, who by the nature of their condition may lack motivation to attend treatment sessions and generally adhere to their treatment plan.

Limitations

A numerical comparison of IECBT clinical outcomes against IAPT's averages in the present study suggests that IECBT is as effective as standard care. The comparison between these two groups is presented to demonstrate general equivalence of IECBT and IAPT services, building on previous results from a clinical trial of IECBT[13] and supporting the effectiveness of this therapy modality in a real-world clinical setting. However, it is important to note there are several limitations for this analysis and caution should be taken not to overinterpret these findings. First, as this was an audit study and not a randomised controlled trial, group comparisons between patients receiving IECBT and IAPT patients are open to the effects of selection bias.

Second, although the IECBT group was matched to the reference group for severity, the aggregated nature of the data published in IAPT's annual reports means that it was not possible to use propensity analyses or selection algorithms to better match the patients who got IECBT to that subset of the patients in IAPT who were most similar to them.

Third, while all patients in the IECBT group received CBT, patients in the IAPT reference group received a range of different therapy types, including IECBT. IECBT is not suitable for all patients, including those at risk and those who are not literate, not fluent English speakers or who do not have access to an internet-connected device. It can be hypothesised that differences in therapy type, together with potential cohort differences in other uncontrolled variables such as presence of secondary comorbid mental health conditions, IQ and socioeconomic status, may also account for variance in clinical outcomes. Furthermore, while there is a large evidence base supporting the efficacy of CBT in general, the issue of side-effects and potential increase in suicidality as a result of therapy remains largely unexplored, for both face-to-face and online CBT modalities. The data collected in the present study and published in IAPT's annual reports did not allow us to explore possible side-effects of IECBT and other types of therapy offered under the IAPT programme.

Regarding predictive analyses in the present study, although these comprehensively included patient and service variables, it should be noted that therapy variables were not included. One advantage of IECBT's unique method, is that it provides full therapy session transcripts. Future work could make use of these data, together with advanced analysis techniques, such as natural language processing and machine learning, to identify therapeutic interventions that lead to the best proximal and distal clinical outcomes in patients.

Implications and future research

A positive aspect of the in-service data collection method used here and in other IAPT services is that replication of these findings is possible in a way that is often cost-prohibitive in clinical trials. Analysis of subsequent cohorts can inform whether these observations generalise to other cohorts and also add to the scientific knowledge of effective CBT change mechanisms. Understanding predictors of good clinical outcomes may facilitate development of improved, patient-focused, stratified/stepped-care allocation models and also enable the development of enhanced therapeutic protocols. Data derived from an outcomes-measurement framework is also of potential value to providers, who can adapt their services to better meet the needs of their patients and consistently monitor service quality and encourage accountability.

In England, continuous monitoring of clinical effectiveness using an outcomes-measurement framework has enabled systematic improvements in the quality and consistency of care delivery and enabled a transition from fee-for-service to fee-for-value payment models.[34] In the USA, as the Centres for Medicare and Medicaid Services begin to implement value-based payment models the importance of understanding what and why treatments work, and what their clinical and economic impact is, becomes evident. Translating the IAPT model, including digital approaches,

not only into the USA but worldwide could have a dual advantage, improving the quality and accountability of mental healthcare while reducing cost by enabling a shift towards capitated and fee-for-value payment models.

Ana Catarino, PhD, Senior Scientist, Clinical Science Laboratory at Ieso, Ieso Digital Health, UK; **Sarah Bateup**, BSc MEd, Chief Clinical Officer, Clinical Science Laboratory at Ieso, Ieso Digital Health, UK; **Valentin Tablan**, PhD, Senior Vice President for Artifical Intelligence, Clinical Science Laboratory at Ieso, Ieso Digital Health, UK; **Katherine Innes**, BSc, Clinical Informatics & Visualisation Specialist, Clinical Science Laboratory at Ieso, Ieso Digital Health, UK; **Stephen Freer**, BA BSc PGDip, UK Clinical Lead, Clinical Science Laboratory at Ieso, Ieso Digital Health, UK; **Andy Richards**, PhD, Director, Ieso Digital Health, UK; **Richard Stott**, DClinPsy, Visiting Associate, Department of Psychology, Institute of Psychiatry, Psychology & Neuroscience, King's College London, UK; **Steven D. Hollon**, PhD, Professor of Psychology, Department of Psychology, Vanderbilt University, USA; **Samuel Robin Chamberlain**, PhD, MRCPsych, Wellcome Trust Clinical Fellow and Honorary Consultant, Department of Psychiatry, University of Cambridge and Cambridge and Peterborough NHS Foundation Trust, UK; **Ann Hayes**, PhD, Independent Pharmaceutical and Digital Health Consultant, Ieso Digital Health, UK; **Andrew D. Blackwell**, PhD, Chief Scientific Officer, Clinical Science Laboratory at Ieso, Ieso Digital Health, UK

Correspondence: Ana Catarino, Clinical Science Laboratory at Ieso, Ieso Digital Health, The Jeffreys Building, Cowley Road, Cambridge CB4 0DS, UK. Email: a.catarino@iesohealth.com

First received 17 Apr 2018, final revision 17 Aug 2018, accepted 5 Sep 2018

Supplementary material

Supplementary material is available online at https://doi.org/10.1192/bjo.2018.57

Funding

This study was funded by Ieso Digital Health. S.R.C. is funded by a Wellcome Trust Clinical Fellowship (Ref: 110049/Z/15/Z).

References

1 Ferrari AJ, Charlson FJ, Norman RE, Patten SB, Freedman G, Murray CJL, et al. Burden of depressive disorders by country, sex, age, and year: findings from the global burden of disease study 2010. *PLoS Med* 2013; **10**: e1001547.

2 McManus S, Bebbington P, Jenkins R, Brugha T. *Mental Health and Wellbeing in England. Health and Social Care Information Centre*, 2016.

3 McCrone P, Dhanasiri S, Patel A, Knapp M, Lawton-Smith S. Paying the price: the cost of mental health care in England to 2026. *Br J Psychiatry* 2008; **184**: 386–92.

4 Clark DM. Implementing NICE guidelines for the psychological treatment of depression and anxiety disorders: the IAPT experience. *Int Rev Psychiatry* 2011; **23**: 318–27.

5 Health and Social Care Information Centre. *Psychological Therapies: Annual Report on the Use of IAPT Services*. Health and Social Care Information Centre, 2016.

6 Olthuis JV, Watt MC, Bailey K, Hayden JA. Therapist-supported Internet cognitive behavioural therapy for anxiety disorders in adults. *Cochrane Database Syst Rev* 2015; **3**: CD011565.

7 Saddichha S, Al-Desouki M, Lamia A, Linden IA, Krausz M. Online interventions for depression and anxiety–a systematic review. *Heal Psychol Behav Med* 2014; **2**: 841–81.

8 Spek V, Cuijpers P, Nyklicek I, Riper H, Keyzer J, Pop V. Internet-based cognitive behaviour therapy for symptoms of depression and anxiety: a meta-analysis. *Psychol Med* 2007; **37**: 319–28.

9 Gega L, Marks I, Mataix-Cols D. Computer-aided CBT self-help for anxiety and depressive disorders: experience of a London clinic and future directions. *J Clin Psychol* 2004; **60**: 147–57.

10 Beattie A, Shaw A, Kaur S, Kessler D. Primary-care patients' expectations and experiences of online cognitive behavioural therapy for depression: a qualitative study. *Health Expect* 2009; **12**: 45–59.

11 Andersson G, Cuijpers P. Pros and cons of online cognitive-behavioural therapy. *Br J Psychiatry* 2008; **193**: 270–1.

12 Dryman MT, McTeague LM, Olino TM, Heimberg RG. Evaluation of an open-access CBT-based Internet program for social anxiety: patterns of use, retention, and outcomes. *J Consult Clin Psychol* 2017; **85**: 988–99.

13 Kessler D, Lewis G, Kaur S, Wiles N, King M, Weich S, et al. Therapist-delivered internet psychotherapy for depression in primary care: a randomised controlled trial. *Lancet* 2009; **374**: 628–34.

14 National Institute for Health and Clinical Excellence. *Common Mental Health Disorders: the NICE Guideline on Identification and Pathways to Care*. The British Psychological Society and The Royal College of Psychiatrists, 2011.

15 El AS, Ljótsson B, Hedman E, Kaldo V, Andersson E, Rück C, et al. Predictors of symptomatic change and adherence in internet-based cognitive behaviour therapy for social anxiety disorder in routine psychiatric care. *PLoS One* 2015; **10**: e0124258.

16 Alaoui S E, Ljotsson B, Hedman E, Svanborg C, Kaldo V, Lindefors N. Predicting outcome in internet-based cognitive behaviour therapy for major depression: a large cohort study of adult patients in routine psychiatric care. *PLoS One* 2016; **11**: e0161191.

17 Roth AD, Pilling S. Accelerated publication using an evidence-based methodology to identify the competences required to deliver effective cognitive and behavioural therapy for depression and anxiety disorders. *Behav Cogn Psychother* 2008; **36**: 129–47.

18 Health and Social Care Information Centre. *IAPT Annual Report for 2015/2016*. NHS Digital, 2017 (http://content.digital.nhs.uk/iaptreports).

19 Gyani A, Shafran R, Layard R, Clark DM. Enhancing recovery rates: lessons from year one of IAPT. *Behav Res Ther* 2013; **51**: 597–606.

20 Kroenke K, Spitzer RL, Williams JB. The PHQ-9: validity of a brief depression severity measure. *J Gen Intern Med* 2001; **16**: 606–13.

21 Spitzer RL, Kroenke K, Williams JBW, Löwe B. A brief measure for assessing generalized anxiety disorder: the GAD-7. *Arch Intern Med* 2006; **166**: 1092–7.

22 Kroenke K, Spitzer RL, Williams JBW, Monahan PO, Lowe B. Anxiety disorders in primary care: prevalence, impairment, comorbidity, and detection. *Ann Intern Med* 2007; **146**: 317–25.

23 R Development Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation, 2016 (http://www.r-project.org/).

24 Martino L, Miguez J. Generalized rejection sampling schemes and applications in signal processing. *Signal Process* 2010; **90**: 2981–95.

25 Myers R. *Classical and Modern Regression with Application (2nd edn)*. Duxbury/Thomson Learning, 1990.

26 Hiller W, Schindler AC, Lambert MJ. Defining response and remission in psychotherapy research: a comparison of the RCI and the method of percent improvement. *Psychother Res* 2012; **22**: 1–11.

27 Lin M, Lucas HC. Too big to fail: large samples and the p-value problem. *Inf Syst Res* 2013; **24**: 938–9.

28 Chen H, Cohen P, Chen S. How big is a big odds ratio? Interpreting the magnitudes of odds ratios in epidemiological studies. *Commun Stat Simul Comput* 2010; **39**: 860–4.

29 Tryon WW, Lewis C. An inferential confidence interval method of establishing statistical equivalence that corrects Tryon's (2001) reduction factor. *Psychol Methods* 2008; **13**: 272–7.

30 Gould RL, Coulson MC, Howard RJ. Efficacy of cognitive behavioral therapy for anxiety disorders in older people: a meta-analysis and meta-regression of randomized controlled trials. *J Am Geriatr Soc* 2012; **60**: 218–29.

31 Thorpe K, Jain S, Joski P. Prevalence and spending associated with patients who have a behavioral health disorder and other conditions. *Health Aff* 2017; **36**: 124–32.

32 Fenton WS, Stover ES. Mood disorders: Cardiovascular and diabetes comorbidity. *Curr Opin Psychiatry* 2006; **19**: 421–7.

33 Clark DM, Canvin L, Green J, Layard R, Pilling S, Janecka M. Transparency about the outcomes of mental health services (IAPT approach): an analysis of public data. *Lancet* 2018; **391**: 679–86.

34 NHS England and NHS Improvement. *Developing an Outcomes-Based Payment Approach for IAPT Services*. NHS Improvement, 2017 (https://improvement.nhs.uk/documents/661/IAPT_Payment_Guidance.pdf).

Letter to the Editor

True technology-enabled mental health care: trust, agency and ageing

Sarah Burch[1], Claire Preston[1], Sarah Bateup[2]

[1]Anglia Ruskin University, Cambridge, UK; [2]Ieso Digital Health, Cambridge, UK

Correspondence to: Dr. Sarah Burch. Anglia Ruskin University, Faculty of Health, Social Care and Education, Young Street Site, East Road, Cambridge, CB1 1PT, UK. Email: sarah.burch@anglia.ac.uk.

Provenance: This is an invited article commissioned by Editor-in-Chief Dr. Steven Tucker, MD (Tucker Medical, Singapore).

Response to: Williams A. True technology-enabled mental health care: when will we trust the computers and robots? Mhealth 2018;4:10.

Received: 21 June 2018; Accepted: 02 August 2018; Published: 07 August 2018.

doi: 10.21037/mhealth.2018.08.02

View this article at: http://dx.doi.org/10.21037/mhealth.2018.08.02

Technology and ageing is a subject which attracts intense interest. In part, this is undoubtedly due to the search for a solution which will help manage the scale, costs and predicted impact of demographic change (1). However, it arguably also reflects a fascination with imagining futures, both in personal and social terms. This may be why there tends to be a focus on the more exotic technological applications. The current emphasis on Ambient and Assisted Living (AAL) contains many pragmatic developments, but can also be used to conjure up exciting visions of older people inhabiting a futuristic domestic landscape staffed by robots. In practice, it is important that this prospect is not allowed to overwrite the more prosaic yet hugely beneficial gains which technology can bring to people who may require increasing levels of support or who struggle to access current-day services.

As noted by Williams (2), to date mental health has not attracted the same enthusiasm for technological innovation as other areas of life. This may stem from the way in which older people's mental health problems are characterised as inevitable, intractable or invisible, in spite of numerous studies showing that there can be good responses to treatment across a range of conditions (3,4). Amongst the array of computer-aided mental health interventions available, internet-enabled cognitive behaviour therapy (IECBT) stands out in that it is not primarily designed to save resources or limit therapist input. It delivers therapist-led sessions in real time which are facilitated through an online portal utilizing written communication (see www.iesohealth.com). Kessler *et al.* (5) have previously demonstrated its effectiveness as a mode of treatment, whilst our analysis of secondary data suggests it has considerable potential as an intervention for older people (6).

There are several ways in which IECBT may be able to make an important contribution to the care and support available to older people. One message from our secondary data analysis was that older people who self-referred for IECBT were assessed as experiencing a wider range of conditions than those who were referred by medical professionals such as GPs (6). This raises some interesting considerations. In terms of how referrals operate, individual clinical commissioning groups (CCGs) in England decide whether to commission IECBT as a service. Patients who reside in areas where IECBT is commissioned may either self-refer or be referred by a medical professional. Once referral or self-referral takes place, the person is assessed by a therapist and the primary diagnosis recorded. The fact that people referred by medical professionals have a narrower range of resultant diagnoses than those who self-refer warrants further exploration. However, a plausible interpretation is that medical professionals either miss certain mental health problems or regard them as unlikely to respond to treatment, hence they only refer a narrow group of older people. Given that studies have shown that older people can be both under-diagnosed and under-treated in relation to conditions like depression (7), this suggests that facilitating self-referral may be one way in which to counter this tendency within services.

Increasing the agency and autonomy of older people may be facilitated in other ways. Allowing self-referral for treatment does not have to be confined to IECBT as opposed to more traditional modes of talking therapies. Yet

one distinctive component of IECBT can be particularly valuable to older people, namely that it provides a permanent written record of the therapy session which has been conducted through text-based communication. This gives the older person the ability to review sessions, which not only reinforces an intrinsic element of the cognitive behavioural therapeutic approach, but also facilitates ownership of the process along with supporting recall of the discussion. Thus technology reinforces the therapeutic exchange.

It is often assumed that communication facilitated by technology will be somehow inferior to all face to face contact. This fear is particularly acute when it comes to talking therapies, as there is such a premium placed on the centrality of the therapeutic relationship. However, it is a misapprehension that relationships can only thrive in person (8,9). Whilst it may not be welcomed by everyone, others can find both anonymity and solace in a relationship conducted at one remove. The computer screen echoes the screen of the confessional, providing a layer of privacy which affords freedom to speak of difficult emotions. Mental health problems are still perceived as sufficiently stigmatizing for this to be a positive boon of technology.

Privacy at a distance can also be accompanied by accessibility (9). One of the biggest difficulties for older people can be physically getting to services even once referred. IECBT means that there is no need to arrange transport or negotiate unfamiliar locations or buildings, as therapy can be accessed at home via the computer. Ageing in place is a common aspiration for older people, yet places can be outgrown or become increasingly unsuitable for their occupants. For ageing in place to be meaningful, the place itself may need to be transformed.

Computer-facilitated therapy has, as previously noted, usually been seen as a cheap option. Most forms are variants on guided self-help, which has the appeal of being far less resource-intensive than conventional therapy. This is not the case with IECBT, as it involves as much therapist input as face to face methods. It is a fallacy that technology will always deliver cost-savings, just as it turned out to be a fallacy that community care would always be cheaper than institutional care, at least not if it were to be delivered to a high standard. Nevertheless, there is still the potential for a reduction in the use of resources, however the resources conserved are those of the older people themselves rather than the NHS. Someone who can access therapy from home does not have to arrange transport, ask family members for help getting to the treatment setting or use up their valuable energy on getting themselves out of the house for another medical appointment rather than a chosen activity. Ageing can involve a dwindling of many kinds of personal and financial resources. If technology enables people to receive therapy at home, these resources can be preserved and even boosted by better mental health. IECBT provides the opportunity to enhance the quality of care provided so that more people benefit from improved mental health. Computer assisted clinical decision support can be used with IECBT to provide information and guidance to therapists whilst they are providing a therapy session. For example, IECBT has had promising results in the lab using interventions that support therapists to provide the right diagnosis and treatment with the aim of enabling more patients to get better.

Of course, IECBT will not suit or work for everyone, whatever age they are. Our research (6) indicates that only a minority of older people access IECBT. Many will not be technologically adept, will find typing slow or difficult, or will value the social contact of face to face interaction. However, if Improving Access to Psychological Therapies (IAPT) as an initiative is to succeed in its goal of increasing uptake amongst older people (10), making effective use of technology is an important strategy to embrace. Broadening the ability to self-refer will permit the increasing numbers of older people who feel comfortable with communicating online to access services which can be transformative in their ability to relieve mental distress. This is an exciting area of study which merits further research.

Acknowledgements

None.

Footnote

Conflicts of Interest: The authors have no conflicts of interest to declare.

References

1. Schulz R, Wahl HW, Matthews JT, et al. Advancing the Aging and Technology Agenda in Gerontology. Gerontologist 2015;55:724-34.
2. Williams A. True technology-enabled mental health care: when will we trust the computers and robots? Mhealth 2018;4:10.
3. Laidlaw K, Kishita N. Age-appropriate augmented

cognitive behaviour therapy to enhance treatment outcomes for late-life depression and anxiety disorders. GeroPsych 2015;28:57-66.
4. Cuijpers P, van Straten A, Smit F, et al. Is psychotherapy for depression equally effective in younger and older adults? A meta-regression analysis. Int Psychogeriatr 2009;21:16-24.
5. Kessler D, Lewis G, Kaur S, et al. Therapist-delivered Internet psychotherapy for depression in primary care: a randomised controlled trial. Lancet 2009;374:628-34.
6. Burch S, Preston C, Bateup S et al. The use of internet-enabled cognitive behavioural therapy in the treatment of depression and anxiety amongst older people. Int J Aging Soc 2017;8:1-11.
7. Overend K, Bosanquet K, Bailey D, et al. Revealing hidden depression in older people: a qualitative study within a randomised controlled trial. BMC Fam Pract 2015;16:142.
8. Suler J. The online disinhibition effect. Cyberpsychol Behav 2004;7:321-6.
9. Preston C, Moore S. Ringing the changes: the role of telephone communication in a helpline and befriending service targeting loneliness in older people. Ageing Soc 2018; doi:10.1017/S0144686X18000120
10. Prina AM, Marioni RE, Hammond GC, et al. Improving access to psychological therapies and older people: findings from the Eastern Region. Behav Res Ther 2014;56:75-81.

doi: 10.21037/mhealth.2018.08.02
Cite this article as: Burch S, Preston C, Bateup S. True technology-enabled mental health care: trust, agency and ageing. mHealth 2018;4:30.

Appendix 4: Burch, Preston, Bateup and Hina (2017)

[Redacted in this version due to copyright. Available at the publisher's website: https://doi.org/10.18848/2576-5310/CGP/v08i01/1-11]

Appendix 5: Ieso Digital Health Plain Language Information for Patients

Why does health and care research use information from patients?
Researchers use data to establish if one treatment method is better or worse than another, or to see whether there are links between recovery rates, treatment methods, and patients' personal circumstances.

Different people can respond to therapy in different ways so by collecting information from lots of people, researchers can work out ways to get more people better.

How does Ieso use my data for research?

When you sign up with us, we will collect information about you for two reasons:

1) The NHS requires that we collect information about our patient care so that they can monitor the quality of the service we provide. This information includes how well patients respond to treatment along with some personal statistics including gender, age and ethnicity. (**see IAPT Minimum Data Set [link]**)
2) To conduct research to understand what works in CBT and further improve treatment. Collecting information from the conversations that happen between patients and therapists helps us to build a greater understanding of the causes of mental illness and what makes treatment work. Click here to find out more. https://www.iesohealth.com/en-gb/data-science
3) We have processes in place to safeguard your privacy. This means that the most identifiable information, such as your name or address, is kept separate from the data we use for research.

How does Ieso use my communications with a therapist?

Being able to look at and analyse conversations between patients and therapists provides us with a unique opportunity to learn how therapy works and improve it.

We have treated over 30,000 patients and we use the data from this for research. We are training computer algorithms to find patterns in the therapy process – so we can see what aspects of therapy are most effective. An example of our research using this approach can be seen here [**link to Ewbank et al., 2019**].

Will a researcher read any of my conversations with my therapist?

They might do, but very rarely. However, every conversation has directly identifiable information removed from it. This means the researcher reading it would never know your name or see any names or specific locations that you or your therapist mentions. The information that links directly to you is removed (see below).

The reason why a research scientist might need to read through a therapy session, is that it is sometimes necessary for a small number of them to be manually "coded". For example, we may need to tag a transcript to mark each time a therapist greets a patient, sets homework, or begins a specific therapy activity.

Will my data be shared outside Ieso in a research context?

Sometimes, but with strict safeguards and we do not share conversations between a patient and therapist.
In order to conduct research to improve treatment, we sometimes partner with researchers outside of Ieso, e.g. university research groups.

When this happens, we remove directly identifiable information (including any names and locations) from the data we share with them, so they will be unable to identify anyone personally.

All partners also sign a legal agreement that any data they receive is kept confidential and secure.

Will my data be made public?

Never in a way that could identify you.

In order to increase the number of people who recover with online CBT, it is important to share our findings with the research community. We publish our findings in peer reviewed journals and present them at academic conferences.

However, we only report statistics, patterns and our conclusions – we don't include details about individuals' treatment or condition. You can see a list of all our published research here **[link to data science page].**

Still have questions?
We're very happy to help if you have more questions. To get in touch please email us at info@iesohealth.com.

Appendix 6: The CTS-R

Student code ……. Score ……...

COGNITIVE THERAPY SCALE - REVISED (CTS-R)

I.-M. Blackburn, I.A. James, D.L. Milne & F.K. Reichelt

Collaborators:
A.Garland, C. Baker, S.H. Standart & A. Claydon

Newcastle upon Tyne, UK - August 2000

Scorer:..... Date:.....

COGNITIVE THERAPY SCALE - REVISED (CTS-R)

The rating of the scale

The present seven point scale (i.e. a 0-6 Likert scale) extends from (0) where the therapist did not adhere to that aspect of therapy (non-adherence) to (6) where there is adherence and very high skill. Thus the scale assesses both adherence to therapy method and skill of the therapist. To aid with the rating of items of the scale, an outline of the key features of each item is provided at the top of each section. A description of the various rating criteria is given in the right hand margin - see example below in Figure 1. Further details are provided in the accompanying manual.

The examples are intended to be used as useful guidelines only. They are not meant to be used as prescriptive scoring criteria, rather providing both illustrative anchor points and guides.

Adjusting the scale in the presence of patient difficulties

The scale's dimensions were devised for patients assessed as being well/moderately suited for cognitive therapy (Safran & Segal, 1990). As such, adjustments may need to be made when patient difficulties are evident (e.g. excessive avoidance). Indeed, with problematic patients it is sometimes difficult to apply CT methods successfully; that is, with desirable change. In such circumstances the rater needs to assess the therapist's therapeutic skills in the application of the methods. Thus even though the therapist may be unsuccessful at promoting change, credit should be given for demonstrations of appropriate skilful therapy.

Safran, J.D. & Segal, Z.V. (1990) *Interpersonal* processes *in* cognitive *therapy*. New York, Basic Books.

Figure 1: Example of the scoring layout

Key features: this is an operationalised description of the item (see examples within the CTS-R).

Mark with an 'X' on the vertical line, using whole and half numbers, the level to which you think the therapist has fulfilled the core function. The descriptive features on the right are designed to guide your decision.

N.B. When rating, take into consideration the appropriateness of therapeutic interventions for stage of therapy and perceived patient difficulty.

Competence level	Examples
Incompetent	0 absence of feature, or highly inappropriate performance
Novice	1 inappropriate performance, with major problems evident
Advanced beginner	2 evidence of competence, but numerous problems and lack of consistency
Competent	3 competent, but some problems and/or inconsistencies
Proficient	4 good features, but minor problems and/or inconsistencies
Expert	5 very good features, minimal problems and/or inconsistencies
	6 excellent performance, even in the face of patient difficulties

* The present scale has incorporated the Dreyfus system (Dreyfus, 1989) for denoting competence, which is described fully in the manual. Please note that the 'top marks (i.e. near the 'expert' end of the continuum) are reserved for those therapists demonstrating highly effective skills, particularly in the face of difficulties (i.e. highly aggressive or avoidant patients; high levels of emotional discharge from the patients; and various situational factors).

The `**Key Features'** describe the important features that need to be considered when scoring each item. When rating the item, you must first identify whether some of the features are present. You must then consider whether the therapist should be regarded as competent with the features. If the therapist includes most of the key features and uses them appropriately (i.e. misses few relevant opportunities to use them), the therapist should be rated very highly.

The `**Examples'** are only guidelines and should not be regarded as absolute rating criteria.

Scoring Distribution

It is important to remember that the scoring profile for this scale should approximate to a normal distribution (i.e. mid-point 3), with relatively few therapists scoring at the extremes.

Dreyfus, H. L. (1989). The Dreyfus model of skill acquisition. *In* J. Burke (ed.) *Competency based education and training.* London: Falmer Press.

ITEM 1 - AGENDA SETTING & ADHERENCE

Key features: To address adequately topics that have been agreed and set in an appropriate way. This involves the setting of discrete and realistic targets collaboratively.. The format for setting the agenda may vary according to the stage of therapy - see manual.

Three features need to be considered when scoring this item:

(i) presence/absence of an agenda which is explicit, agreed and prioritised, and feasible in the time available;

(ii) appropriateness of the contents of the agenda (to stage of therapy, current concerns etc.), a standing item being a review of the homework set previously;

(iii) appropriate adherence to the agenda.

Mark with an 'X' on the vertical line, the level to which you think the therapist has fulfilled the core function. The descriptive features on the right are designed to guide your decision.

NB: Agenda setting requires collaboration and credit for this should be given here, and here alone. Collaboration occurring at any other phase of the session should be scored under Item 3 (Collaboration).

Competence level	Examples NB: Score according to features, not examples!
0	No agenda set, highly inappropriate agenda set, or agenda not adhered to.
1	Inappropriate agenda set (e.g. lack of focus, unrealistic, no account of patient's presentation, homework not reviewed).
2	An attempt at an agenda made, but major difficulties evident (e.g. unilaterally set). Poor adherence.
3	Appropriate agenda, which was set well, but some difficulties evident (e.g. poor collaboration). Some adherence.
4	Appropriate agenda, minor difficulties evident (e.g. no prioritisation), but appropriate features covered (e.g. review of homework). Moderate adherence.
5	Appropriate agenda set with discrete and prioritised targets - review at the end. Agenda adhered to. Minimal problems.
6	Excellent agenda set, or highly effective agenda set in the face of difficulties.

ITEM 2 - FEEDBACK

Key features: The patient's and therapist's understanding of key issues should be helped through the use of two-way feedback: The two major forms of feeding back information are through general summary and chunking of important units of information. The use of appropriate feedback helps both the therapist to understand the patient's situation, and the patient to synthesise material enabling him/her to gain major insight and make therapeutic shifts. It also helps to keep the patient focused.

Three features need to be considered when scoring this item:

(i) presence and frequency, or absence, of feedback. Feedback should be given/elicited throughout the therapy - with major summaries both at the beginning (review of week) and end (session summary), while topic reviews (i.e. chunking) should occur throughout the session;
(ii) appropriateness of the contents of the feedback;
(iii) manner of its delivery and elicitation (NB: can be written).

Competence level	Examples NB: Score according to features, not examples!
—	Absence of feedback or highly inappropriate feedback.
0	Minimal appropriate feedback (verbal and/or written).
1	Appropriate feedback, but not given frequently enough by therapist, with insufficient attempts to elicit and give feedback, e.g. feedback too vague to provide opportunities for understanding and change.
2	Appropriate feedback given and elicited frequently, although some difficulties evident in terms of content or method of delivery.
3	Appropriate feedback given and elicited frequently, facilitating moderate therapeutic gains. Minor problems evident (eg. inconsistent).
4	Highly appropriate feedback given and elicited regularly, facilitating shared understanding and enabling significant therapeutic gains. Minimal problems.
5	Excellent use of feedback, or highly effective feedback given and elicited regularly in the face of difficulties.
6	

ITEM 3 - COLLABORATION

Key features: The patient should be encouraged to be active in the session. There must be clear evidence of productive teamwork, with the therapist skillfully encouraging the patient to participate fully (e.g. through questioning techniques, shared problem solving and decision making) and take responsibility. However, the therapist must not allow the patient to ramble in an unstructured way.

Three features need to be considered: the therapist style should encourage effective teamwork through his/her use of:

(i) verbal skills (e.g. non-hectoring);
(ii) non-verbal skills (e.g. attention and use of joint activities);
(iii) sharing of written summaries.

NB: Questioning is a central feature with regard to this item, but questions designed to facilitate reflections and self discovery should be scored under Item 9 (Guided Discovery).

Competence level	Examples NB: Score according to features, not examples!
0	Patient is actively prevented or discouraged from being collaborative.
1	The therapist is too controlling, dominating, or passive.
2	Some occasional attempt at collaboration, but didactic style or passivity of therapist encourages passivity or other problems in the therapeutic relationship.
3	Teamwork evident, but some problems with collaborative set (e.g. not enough time allowed for the patient to reflect and participate actively).
4	Effective teamwork is evident, but not consistent. Minor problems evident.
5	Effective teamwork evident throughout most of the session, both in terms of verbal content and use of written summaries. Minimal problems.
6	Excellent teamwork, or highly effective teamwork in the face of difficulties.

ITEM 4 - PACING AND EFFICIENT USE OF TIME

Key features: The session should be well 'time managed' in relation to the agenda, with the session flowing smoothly through discrete start, middle, and concluding phases. The work must be paced well in relation to the patient's needs, and while important issues need to be followed, unproductive digressions should be dealt with smoothly. The session should not go over time, without good reason.

Three features need to be considered:

(i) the degree to which the session flows smoothly through the discrete phases;
(ii) the appropriateness of the pacing throughout the session;
(iii) the degree of fit to the learning speed of the patient.

Competence level	Examples NB: Score according to features, not examples!
0	Poor time management leads either to an aimless or overly rigid session.
1	The session is too slow or too fast for the current needs and capacity of the patient.
2	Reasonable pacing, but digression or repetitions from therapist and/or patient lead to inefficient use of time; unbalanced allocation of time, over time.
3	Good pacing evident some of the time, but diffuse at times. Some problems evident.
4	Balanced allocation of time with discrete start, middle and concluding phases evident. Minor problems evident.
5	Good time management skills evident, session running smoothly. Therapist working effectively in controlling the flow within the session. Minimal problems.
6	Excellent time management, or highly effective management evident in the face of difficulties.

ITEM 5 - INTERPERSONAL EFFECTIVENESS

Key features: The patient is put at ease by the therapist's verbal and non-verbal (e.g. listening skills) behaviour. The patient should feel that the core conditions (i.e. warmth, genuineness, empathy and understanding) are present. However, it is important to keep professional boundaries. In situations where the therapist is extremely interpersonally effective, he/she is creative, insightful and inspirational.

Three features need to be considered:

(i) empathy - the therapist is able to understand and enter the patient's feelings imaginatively and uses this understanding to promote change;
(ii) genuineness - the therapist has established a trusting working relationship;
(iii) warmth - the patient seems to feel liked and accepted by the therapist.

Competence level	Examples NB: Score according to features, not examples!
0	Therapist's manner and interventions make the patient disengage and become distrustful and/or hostile (absence of/or excessive i, ii, iii).
1	Difficulty in showing empathy, genuineness and warmth.
2	Therapist's style (e.g. intellectualisation) at times impedes his/her empathic understanding of the patient's communications.
3	The therapist is able to understand explicit meanings of patient's communications, resulting in some trust developing. Some evidence of inconsistencies in sustaining relationship.
4	The therapist is able to understand the implicit, as well as the explicit meanings of the patient's communications and demonstrates it in his/ her manner. Minor problems evident (e.g. inconsistent).
5	The therapist demonstrates very good interpersonal effectiveness. Patient appears confident that he/she is being understood, which facilitates self-disclosure. Minimal problems.
6	Highly interpersonally effective, even in the face of difficulties.

ITEM 6 — ELICITING OF APPROPRIATE EMOTIONAL EXPRESSION

Key features: The therapist facilitates the processing of appropriate levels of emotion by the patient. Emotional levels that are too high or too low are likely to interfere with therapy. The therapist must also be able to deal effectively with emotional issues which interfere with effective change (e.g. hostility, anxiety, excessive anger). Effective facilitation will enable the patient to access and express his/her emotions in a way that facilitates change.

Three features have to be considered:

(i) facilitation of access to a range of emotions;
(ii) appropriate use and containment of emotional expression;
(iii) facilitation of emotional expression; encouraging appropriate access and differentiation of emotions.

Competence Level	Examples NB. Score according to features, not examples!
0	Patient is under- or over stimulated (e,g, his her feelings are ignored or dismissed or allowed to reach an unmanaged pitch). Or the therapist's own mood or strategies (e.e. intellectualization) adversely influences the session.
1	Failure to facilitate access to, and expression of, appropriate emotional expression.
2	Facilitation of appropriate emotional expression evident, but many relevant opportunities missed.
3	Some effective facilitation of appropriate emotional expression, created and/or maintained. Patient enabled to become slightly more aware.
4	Effective facilitation of appropriate emotional expression leading to the patient becoming more aware of relevant emotions. Minor problems evident.
5	Very effective facilitation of emotional expression, optimally arousing the patient's motivation and awareness. Good expression of relevant emotions evident- done in an effective manner. Minimal problems.
6	Excellent facilitation of appropriate emotional expression, or effective facilitation in the face of difficulties

ITEM.7 - ELICITING KEY COGNITIONS

Key features: To help the patient gain access to his/her cognitions (thoughts, assumptions and beliefs) and to understand the relationship between these and their distressing emotions. This can be done through the use of questioning, diaries and monitoring procedures.

Three features need to be considered:

(i) eliciting cognitions that are associated with distressing emotions (i.e. selecting key cognitions or hot thoughts);

(ii) the skillfulness and breadth of the methods used (i.e. Socratic questioning; appropriate monitoring, downward arrowing, imagery, role-plays, etc.);

(iii) choosing the appropriate level of work for the stage of therapy (i.e. automatic thoughts, assumptions, or core beliefs).

NB: This item is concerned with the general work done with eliciting cognitions. If any specific cognitive or behavioural change methods are used, they should be scored under item 11 (change methods).

Competence level	Examples NB: Score according to features, not examples!
0	Therapist fails to elicit relevant cognitions.
1	Inappropriate cognitions and emotions selected, or key cognitions/emotions ignored.
2	Some cognitions/emotions (or one key cognition, e.g. core belief) elicited, but links between cognitions and emotions not made clear to patient.
3	Some cognitions/emotions (or one key cognition) elicited in a competent way, although some problems evident.
4	A number of cognitions and emotions (or one key cognition) elicited in verbal or written form, leading to a new understanding of their relationship. Minor problems evident.
5	Effective eliciting and selection of a number of cognitions/emotions (or one key cognition), which are generally dealt with appropriately. Minimal problems.
6	Excellent work done on key cognition(s) and emotion(s), even in the face of difficulties.

ITEM 8 - ELICITING AND PLANNING BEHAVIOURS

Key features: To help the patient gain insight into the effect of his/her behaviours with respect to the problems. This can be done through the use of questioning; diaries and monitoring procedures. The therapist works with the patient to plan strategies either to overcome or disrupt dysfunctional behavioural patterns.

Two features need to be considered:

(i) eliciting behaviours and plans that are associated with distressing emotions;
(ii) the skilfulness and breadth of the methods used (i.e. Socratic questioning; appropriate monitoring, downward arrowing, imagery, role-plays, etc.);

NB: This item is concerned with the general work done with eliciting behaviours and plans. If any specific cognitive or behavioural change methods are used, they should be scored under item 11 (change methods).

Competence level	Examples NB: Score according to features, not examples!
0	Therapist fails to elicit relevant behaviours and plans.
1	Inappropriate behaviours focused on and/or plans generated.
2	Some behaviours and plans elicited, but links between behaviours, cognitions and emotions not made clear to patient.
3	Some behaviours and plans elicited in a competant way, although some problems evident.
4	A number of behaviours and plans elicited in verbal or written form, leading to a new understanding of their importance in maintaining problems. Minor difficulties evident.
5	Effective eliciting and selection of a number of behaviours and plans, which are generally dealt with appropriately. Minimal problems.
6	Excellent work done on behaviours and plans, even in the face of difficulties.

ITEM 9 - GUIDED DISCOVERY

Key features: The patient should be helped -to -develop hypotheses regarding his/her current situation and to generate potential solutions' for him/herself. The patient is helped to develop a range of perspectives regarding his/her experience. Effective guided discovery will create doubt where previously there was certainty, thus providing the opportunity for re-evaluation and new learning to occur.

Two elements need to be considered:

(i) the style of the therapist - this should be.open and inquisitive;
(ii) the effective use of questioning techniques (e.g. Socratic questions) should encourage the patient to discover useful information that can be used to help him/her to gain a better level of understanding.

Competence level	Examples NB: Score according to features, not examples!
0	No attempt at guided discovery (e.g. hectoring and lecturing).
1	Little opportunity for discovery by patient. Persuasion and debate used excessively.
2	Minimal opportunity for discovery. Some use of questioning, but unhelpful in assisting the patient to gain access to his/her thoughts or emotions or to make connections between themes.
3	Some reflection evident. Therapist uses primarily a questioning style which is following a productive line of discovery.
4	Moderate degree of discovery evident. Therapist uses a questioning style with skill, and this leads to some synthesis. Minor problems evident.
5	Effective reflection evident. Therapist uses skilful questioning style leading to reflection, discovery, and synthesis. Minimal problems.
6	Excellent guided discovery leading to a deep patient understanding. Highly effective discovery produced in the face of difficulties, with evidence of a deeper understanding having been developed.

ITEM 10 - CONCEPTUAL INTEGRATION

Key features: The patient should be helped to gain an appreciation of the history, triggers and maintaining features of his/her problem in order to bring about change in the present and future. The therapist should help the patient to gain an understanding of how his/her perceptions and interpretations, beliefs, attitudes and rules relate to his/her problem. A good conceptualisation will examine previous cognitions and coping strategies as well as current ones. This theory-based understanding should be well integrated and used to guide the therapy forward.

Two features need to be considered:

(i) the presence/absence of an appropriate conceptualisation which is in line with goals of therapy;
(ii) the manner in which the conceptualisation is used (e.g. used as the platform for interventions, homework etc.).

NB: This item is to do with therapeutic integration (using theory to link present, past and future). If the therapist deals specifically with cognitions and emotions, this should be scored under Items 6 (Facilitation of Emotional Expression) and 7 (Eliciting Key of Cognitions)

Competence level	Examples NB: Score according to features, not examples!
0	The absence of an appropriate conceptualisation.
1	The lack, or inappropriateness or misapplication of a conceptualisation leads to a neutral impact (e.g. interferes with progress or leads to aimless application of procedures).
2	Some rudimentary conceptualisation arrived at, but not well integrated with goals of therapy. Does not lead to a clear rationale for interventions.
3	Cognitive conceptualisation partially developed with some integration, but some difficulties evident (e.g. in synthesising and in sharing it with the patient). Leads to coherent interventions.
4	Cognitive conceptualisation is moderately developed and integrated within the therapy. Minor problems evident.
5	Cognitive conceptualisation is very well developed and integrated within the therapy - there is a credible cognitive understanding leading to major therapeutic shifts. Minimal problems.
6	Excellent development and integration evident, or highly effective in the face of difficulties.

ITEM 11- APPLICATION OF CHANGE METHODS

Key features: Therapist skillfully uses, and helps the patient to use, appropriate cognitive and behavioural techniques in line with the formulation. The therapist helps the patient devise appropriate cognitive methods to evaluate the key cognitions associated with distressing emotions, leading to major new perspectives and shifts in emotions. The therapist also helps the patient to apply behavioural techniques in line with the formulation. The therapist helps the patient to identify potential difficulties and think through the cognitive rationales for performing the tasks. The methods provide useful ways for the patient to test-out cognitions practically and gain experience in dealing with high levels of emotion. The methods also allow the therapist to obtain feedback regarding the patient's level of understanding of prospective practical assignments (i.e. by the patient performing the task in- session).

Three features need to be considered:

(i) the appropriateness and range of both cognitive methods (e.g. cognitive change diaries, continua, distancing, responsibility charts, evaluating alternatives, examining pros and cons, determining meanings, imagery restructuring, etc.) and behavioural methods (e.g. behavioural diaries, behavioural tests, role play, graded task assignments, response prevention, reinforcement of patient's work, modeling, applied relaxation, controlled breathing, etc.);

(ii) the skill in the application of the methods - however, skills such as feedback, interpersonal effectiveness, etc. should be rated separately under their appropriate items;

(iii) the suitability of the methods for the needs of the patient (i.e. neither too difficult nor complex).

NB: This item is not concerned with accessing or identifying thoughts, rather with their re-evaluation.

Competence level	Examples NB: Score according to features, not examples!
0	Therapist fails to use or misuses appropriate cognitive and behavioural methods.
1	Therapist applies either insufficient or inappropriate methods, and/or with limited skill or flexibility.
2	Therapist applies appropriate methods, but major difficulties evident.
3	Therapist applies a number of methods in competent ways, although some problems evident (e.g. the interventions are incomplete).
4	Therapist applies a range of methods with skill and flexibility, enabling the patient to develop new perspectives. Minor problems evident.
5	Therapist systematically applies an appropriate range of methods in a creative, resourceful and effective manner. Minimal problems.
6	Excellent range and application, or successful application in the face of difficulties.

ITEM 12 - HOMEWORK SETTING

Key features: This aspect concerns the setting of an appropriate homework task, one with clear and precise goals. The aims should be to negotiate an appropriate task for the stage of therapy in line with the conceptualisation; to ensure the patient understands the rationale for undertaking the task; to test out ideas, try new experiences, predict and deal with potential obstacles, and experiment with new ways of responding.

There are three aspects to this item:

(i) presence/absence of a homework task in which clear and precise goals have been set;
(ii) the task should be derived from material discussed in the session, such that there is a clear understanding of what will be learnt from performing the task;
(iii) the homework task should be set jointly, and sufficient time should be allowed for it to be explained clearly (i.e. explain, discuss relevance, predict obstacles, etc.).

NB: Review of homework from the previous session should be rated in Item 1(Agenda Setting)

Competence level	Examples NB: Score according to features, not examples!
0	Therapist fails to set homework, or sets inappropriate homework.
1	Therapist does not negotiate homework. Insufficient time allotted for adequate explanation, leading to ineffectual task being set.
2	Therapist negotiates homework unilaterally and in a routine fashion, without explaining the rationale for new homework.
3	Therapist has set an appropriate new homework task, but some problems evident (e.g. not explained sufficiently and/or not developed jointly).
4	Appropriate new homework jointly negotiated with a clear goals and rationales. However, minor problems evident.
5	Appropriate homework negotiated jointly and explained well, including an exploration of potential obstacles. Minimal problems.
6	Excellent homework negotiated, or appropriate one set in the face of difficulties.

Appendix 7: The Participant Information Sheet

Participant Information Sheet

New Research Study
Understanding Therapist Variables

You are invited to take part in a research study

My name is Sarah Bateup I am the Chief Clinical Officer at Ieso Digital Health and a Professional Doctorate student at Anglia Ruskin University in Cambridge. As part of my doctorate I am investigating the therapist variables that correlate with clinical outcomes within IAPT. If you have any questions about this study, you can contact me by telephone on 01954 230066 or email at sarah.bateup@pgr.anglia.ac.uk Alternatively, you may contact my supervisor, Dr. Sarah Burch at Anglia Ruskin University by telephone on 01245 493131, or by email at sarah.burch@anglia.ac.uk.

You are free to decide whether or not to take part in this study and if you decide to participate you may change your mind at any time.

Overview of the study

The measurement and evaluation of clinical effectiveness has become routine practice within CBT services in the UK, particularly within IAPT. Therapists are routinely asked to use tools such as the CTS-R for the purposes of continuing professional development and all IAPT services record the clinical outcome data for each of its therapists. Whilst all IAPT services are mandated to report clinical outcomes to NHS England and IAPT has reported that good clinical outcomes correlate with adherence to NICE guidelines, very little is actually known about the therapist variables that correlate with clinical outcome. This is because it has never been possible to observe what therapists are doing in the therapy room in IAPT services. Delivering CBT using the Ieso Method offers, for the first time ever, the opportunity to understand what therapists are doing with their patients.

I am interested to understand which therapist variables correlate with clinical outcome. The findings from this study will help cognitive behavioural therapists learn more about effective clinical practice so that more patients can get better. These findings will also enable supervisors to provide more effective support and training to therapists. By participating in this study, you are enabling others to learn more about how to enhance and improve the practice of CBT.

Why am I being invited to participate in this study?

All BABCP accredited therapists who have had at least 20 patients at Ieso Digital Health are being invited to participate in this study.

What do you need to know about the methods that will be used in this study?

If you decide to take part in this study, you will be asked to complete a short survey via the Ieso Digital Health 'Hub' (online learning management system). This will take no more than 2 minutes of your time. The survey will ask you 5 short questions that relate to your core profession, your training and years of experience. This information will be used to analyse whether there is a correlation between clinical outcomes and a therapist's professional background or years of experience.

If you decide to participate, I will use the clinical data that has already been collected by the clinical team at Ieso Digital Health relating to the patients you have treated. This data relates to the recovery rates of your patients. I will use this data to understand if there is a correlation between clinical outcomes, professional experience and CTS-R scores. As you know, Ieso Digital Health routinely undertakes CTS-Rs of randomly selected therapy sessions and reviews all of its therapists' clinical work. If you consent to participate in the study this data will be used to understand the variables (CTS-R scores, attendance at supervision etc) that correlate with clinical outcome.

If you agree to participate in this study your Clinical Supervisor will be informed. Your supervisor will be asked about how you use the supervision that is provided for you at Ieso Digital Health. Your supervisor will be asked: whether you attend and prepare for supervision, whether you are able to reflect on your work and use any feedback that is provided in order to enhance your clinical skills. Supervision in this context includes supervision in groups, individual supervision by phone or email and supervision provided by the Hub forums (Ieso Digital Health's online learning management system).

The findings from this study will be documented in my thesis, published in professional journals and presented at conferences. You will not be identified at any time. If you agree to participate in this study, you will be given a unique identification number. I am interested in the collective themes drawn from the data of many therapists, not individual therapists.

All data will be held in accordance with the Data Protection Act (1998). In addition, the data that is collected for the purposes of this study will be stored on a secure and encrypted computer that complies with Ieso Digital Health's ISO27001 (data security and storage) certification. Your personal information is strictly confidential and will not be published, shared or discussed with anyone for the purposes of this study.

What are the benefits of participating in this study?

You are unlikely to gain personally from this study. However,

the findings from this study will help cognitive behavioural therapists learn more about effective clinical practice so that more patients can get better. These findings will also enable supervisors to provide more effective support and training to therapists. By participating in this study, you are enabling others to learn more about how to enhance and improve the practice of CBT.

What if I have questions about the data that is collected about me?

Ieso Digital Health routinely reviews the work of all its therapists and uses this data to drive a programme of continuing professional development. This is not undertaken as a punitive process but serves the function of supporting and enabling therapists to help more patients get better. This data can be requested by any therapist at any time. If you would like to see a copy of your data or you have any questions about your data, please contact me by telephone on 01954 230066 or by email at sarah.bateup@pgr.anglia.ac.uk

What if I change my mind about participating in the study?

You can change your mind about participating in this study at any time. If you no longer wish to participate in this study you should contact me, or my supervisor, Dr. Sarah Burch. You can contact me by telephone on 01954 230066 or email at sarah.bateup@pgr.anglia.ac.uk. Alternatively, you may contact my supervisor, Dr. Sarah Burch at Anglia Ruskin University by telephone on 01245 493131, or by email at sarah.burch@anglia.ac.uk.

If you change your mind your data will be removed from the analysis in this study.

Who has reviewed this study?

As you know, all research studies are reviewed by an independent group of people, called a research ethics committee, to protect your safety, rights, wellbeing and dignity. The study has received ethical approval from the Departmental Research Ethics Panel in the Faculty of Health, Social Care and Education at Anglia Ruskin University, Cambridge.

In addition, Ieso Digital Health has given their permission for me to use the data held by them for the purposes of this research study.

Complaints

If you have any complaints about this study please speak to me, or my supervisor, in the first instance. You can contact me by telephone on 01954 230066 or email at sarah.bateup@pgr.anglia.ac.uk My supervisor is Dr. Sarah Burch and you can contact her at Anglia Ruskin University by telephone on 01245 493131, or by email at sarah.burch@anglia.ac.uk.
However, if you wish to make a formal complaint, the details below will help you.
Email address: complaints@anglia.ac.uk

Postal address: Office of the Secretary and Clerk, Anglia Ruskin University, Bishops Hall Lane, Chelmsford, Essex, CM12 1SQ.

Participant Information Sheet 10/01/2018 version 4

Appendix 8: SPSS output from hierarchical loglinear analysis

```
NEW FILE.
DATASET NAME DataSet3 WINDOW=FRONT.
PRESERVE.
SET DECIMAL DOT.

GET DATA  /TYPE=TXT
  /FILE="/Users/sbateup/Desktop/F.csv"
  /ENCODING='UTF8'
  /DELCASE=LINE
  /DELIMITERS=""
  /ARRANGEMENT=DELIMITED
  /FIRSTCASE=2
  /DATATYPEMIN PERCENTAGE=95.0
  /VARIABLES=
  F AUTO
  /MAP.
RESTORE.

CACHE.
EXECUTE.

Data written to the working file.
1 variables and 626 cases written.
Variable: F              Type: Number  Format : PCT6.3     One or more values were set to
system-missing.

DATASET NAME DataSet4 WINDOW=FRONT.
DESCRIPTIVES VARIABLES=F
  /STATISTICS=MEAN STDDEV MIN MAX KURTOSIS.
Descriptives
Notes
Output Created                    15-JAN-2019 18:36:12
Comments
Input    Data     /Users/sbateup/Desktop/F.csv
         Active Dataset DataSet4
         Filter    <none>
         Weight<none>
         Split File          <none>
         N of Rows in Working Data File          626
Missing Value Handling          Definition of Missing   User defined missing values are treated
as missing.
         Cases Used      All non-missing data are used.
Syntax            DESCRIPTIVES VARIABLES=F
  /STATISTICS=MEAN STDDEV MIN MAX KURTOSIS.
Resources       Processor Time          00:00:00.00
```

Elapsed Time 00:00:00.00

[DataSet4]

Descriptive Statistics

	N Statistic	Minimum Statistic	Maximum Statistic	Mean Statistic	Std. Deviation Statistic	Kurtosis Statistic	Statistic Std. Error
F	624	5.000%	81.900%	40.17615%	13.151269%	-.043	.195
Valid N (listwise)	624						

```
FREQUENCIES VARIABLES=F
  /NTILES=4
  /NTILES=4
  /STATISTICS=STDDEV MEAN MEDIAN MODE SKEWNESS SESKEW KURTOSIS SEKURT
  /HISTOGRAM NORMAL
  /ORDER=ANALYSIS.
```

Frequencies

Notes

Output Created		15-JAN-2019 18:37:42
Comments		
Input	Data	/Users/sbateup/Desktop/F.csv
	Active Dataset	DataSet4
	Filter	<none>
	Weight	<none>
	Split File	<none>
	N of Rows in Working Data File	626
Missing Value Handling	Definition of Missing	User-defined missing values are treated as missing.
	Cases Used	Statistics are based on all cases with valid data.
Syntax	FREQUENCIES VARIABLES=F /NTILES=4 /NTILES=4 /STATISTICS=STDDEV MEAN MEDIAN MODE SKEWNESS SESKEW KURTOSIS SEKURT /HISTOGRAM NORMAL /ORDER=ANALYSIS.	
Resources	Processor Time	00:00:01.81
	Elapsed Time	00:00:01.00

Statistics

F

N	Valid	624

	Missing	2
Mean		40.17615%
Median		40.28000%
Mode		50.000%
Std. Deviation		13.151269%
Skewness		.201
Std. Error of Skewness		.098
Kurtosis		-.043
Std. Error of Kurtosis		.195
Percentiles	25	30.60000%
	50	40.28000%
	75	48.61000%

F

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	5.000%1	.2	.2	.2	
	6.000%1	.2	.2	.3	
	8.000%1	.2	.2	.5	
	11.000%	1	.2	.2	.6
	12.500%	2	.3	.3	1.0
	13.890%	1	.2	.2	1.1
	15.000%	4	.6	.6	1.8
	15.280%	3	.5	.5	2.2
	16.600%	1	.2	.2	2.4
	16.670%	1	.2	.2	2.6
	16.700%	2	.3	.3	2.9
	17.000%	1	.2	.2	3.0
	18.000%	6	1.0	1.0	4.0
	18.060%	2	.3	.3	4.3
	18.100%	2	.3	.3	4.6
	19.000%	4	.6	.6	5.3
	19.400%	2	.3	.3	5.6
	19.440%	4	.6	.6	6.3
	20.000%	2	.3	.3	6.6
	20.800%	2	.3	.3	6.9
	20.830%	2	.3	.3	7.2
	21.000%	5	.8	.8	8.0
	22.000%	8	1.3	1.3	9.3
	22.200%	3	.5	.5	9.8
	22.220%	2	.3	.3	10.1
	22.920%	1	.2	.2	10.3
	23.600%	2	.3	.3	10.6
	23.610%	3	.5	.5	11.1
	24.000%	3	.5	.5	11.5
	25.000%	14	2.2	2.2	13.8

25.700%	1	.2	.2	13.9
26.000%	11	1.8	1.8	15.7
26.390%	6	1.0	1.0	16.7
26.400%	1	.2	.2	16.8
27.000%	3	.5	.5	17.3
27.780%	2	.3	.3	17.6
27.800%	4	.6	.6	18.3
28.000%	6	1.0	1.0	19.2
28.470%	2	.3	.3	19.6
28.500%	1	.2	.2	19.7
29.000%	9	1.4	1.4	21.2
29.170%	4	.6	.6	21.8
29.200%	3	.5	.5	22.3
30.000%	10	1.6	1.6	23.9
30.560%	5	.8	.8	24.7
30.600%	3	.5	.5	25.2
31.000%	7	1.1	1.1	26.3
31.300%	1	.2	.2	26.4
31.900%	2	.3	.3	26.8
31.940%	5	.8	.8	27.6
32.000%	12	1.9	1.9	29.5
33.000%	9	1.4	1.4	30.9
33.300%	4	.6	.6	31.6
33.330%	5	.8	.8	32.4
34.000%	6	1.0	1.0	33.3
34.700%	5	.8	.8	34.1
34.720%	5	.8	.8	34.9
35.000%	7	1.1	1.1	36.1
35.400%	1	.2	.2	36.2
36.000%	13	2.1	2.1	38.3
36.100%	7	1.1	1.1	39.4
36.110%	9	1.4	1.4	40.9
36.800%	1	.2	.2	41.0
37.000%	1	.2	.2	41.2
37.150%	1	.2	.2	41.3
37.500%	10	1.6	1.6	42.9
38.000%	15	2.4	2.4	45.4
38.190%	1	.2	.2	45.5
38.890%	7	1.1	1.1	46.6
38.900%	2	.3	.3	47.0
39.000%	7	1.1	1.1	48.1
40.000%	11	1.8	1.8	49.8
40.280%	6	1.0	1.0	50.8
40.300%	4	.6	.6	51.4
40.970%	2	.3	.3	51.8
41.000%	4	.6	.6	52.4
41.200%	1	.2	.2	52.6

41.670%	9	1.4	1.4	54.0
41.700%	8	1.3	1.3	55.3
42.000%	6	1.0	1.0	56.3
43.000%	10	1.6	1.6	57.9
43.060%	4	.6	.6	58.5
43.100%	10	1.6	1.6	60.1
43.750%	2	.3	.3	60.4
43.800%	1	.2	.2	60.6
44.000%	9	1.4	1.4	62.0
44.400%	4	.6	.6	62.7
44.440%	4	.6	.6	63.3
45.000%	3	.5	.5	63.8
45.100%	2	.3	.3	64.1
45.140%	1	.2	.2	64.3
45.800%	4	.6	.6	64.9
45.830%	10	1.6	1.6	66.5
46.000%	8	1.3	1.3	67.8
46.500%	1	.2	.2	67.9
46.530%	4	.6	.6	68.6
47.000%	13	2.1	2.1	70.7
47.200%	5	.8	.8	71.5
47.220%	4	.6	.6	72.1
47.900%	1	.2	.2	72.3
47.920%	2	.3	.3	72.6
48.000%	2	.3	.3	72.9
48.600%	5	.8	.8	73.7
48.610%	11	1.8	1.8	75.5
49.000%	6	1.0	1.0	76.4
49.300%	2	.3	.3	76.8
49.310%	2	.3	.3	77.1
50.000%	18	2.9	2.9	80.0
50.690%	1	.2	.2	80.1
51.000%	6	1.0	1.0	81.1
51.390%	9	1.4	1.4	82.5
51.400%	5	.8	.8	83.3
52.000%	1	.2	.2	83.5
52.080%	4	.6	.6	84.1
52.100%	1	.2	.2	84.3
52.780%	2	.3	.3	84.6
52.800%	2	.3	.3	84.9
53.000%	2	.3	.3	85.3
54.000%	6	1.0	1.0	86.2
54.170%	5	.8	.8	87.0
54.200%	4	.6	.6	87.7
54.860%	1	.2	.2	87.8
55.000%	1	.2	.2	88.0
55.560%	1	.2	.2	88.1

	55.600%	2	.3	.3	88.5
	56.000%	3	.5	.5	88.9
	56.300%	1	.2	.2	89.1
	56.900%	4	.6	.6	89.7
	56.940%	4	.6	.6	90.4
	57.000%	3	.5	.5	90.9
	58.000%	6	1.0	1.0	91.8
	58.300%	3	.5	.5	92.3
	58.330%	5	.8	.8	93.1
	59.700%	2	.3	.3	93.4
	59.720%	1	.2	.2	93.6
	60.000%	3	.5	.5	94.1
	61.000%	1	.2	.2	94.2
	61.100%	1	.2	.2	94.4
	61.110%	1	.2	.2	94.6
	62.500%	3	.5	.5	95.0
	63.000%	1	.2	.2	95.2
	63.890%	1	.2	.2	95.4
	63.900%	1	.2	.2	95.5
	64.000%	4	.6	.6	96.2
	65.000%	2	.3	.3	96.5
	65.300%	1	.2	.2	96.6
	66.700%	1	.2	.2	96.8
	68.000%	1	.2	.2	97.0
	68.100%	3	.5	.5	97.4
	69.400%	4	.6	.6	98.1
	70.800%	1	.2	.2	98.2
	71.000%	1	.2	.2	98.4
	72.000%	1	.2	.2	98.6
	72.200%	1	.2	.2	98.7
	72.220%	1	.2	.2	98.9
	73.600%	1	.2	.2	99.0
	74.000%	3	.5	.5	99.5
	77.800%	1	.2	.2	99.7
	81.000%	1	.2	.2	99.8
	81.900%	1	.2	.2	100.0
	Total	624	99.7	100.0	
Missing	System	2	.3		
Total		626	100.0		

```
EXAMINE VARIABLES=F
  /PLOT BOXPLOT STEMLEAF HISTOGRAM
  /COMPARE GROUPS
  /PERCENTILES(5,10,25,50,75,90,95) HAVERAGE
  /STATISTICS DESCRIPTIVES
```

/CINTERVAL 95
/MISSING LISTWISE
/NOTOTAL.

Explore

Notes

Output Created	15-JAN-2019 18:51:55
Comments	
Input Data	/Users/sbateup/Desktop/F.csv
Active Dataset	DataSet4
Filter	<none>
Weight	<none>
Split File	<none>
N of Rows in Working Data File	626
Missing Value Handling Definition of Missing	User-defined missing values for dependent variables are treated as missing.
Cases Used	Statistics are based on cases with no missing values for any dependent variable or factor used.
Syntax	EXAMINE VARIABLES=F /PLOT BOXPLOT STEMLEAF HISTOGRAM /COMPARE GROUPS /PERCENTILES(5,10,25,50,75,90,95) HAVERAGE /STATISTICS DESCRIPTIVES /CINTERVAL 95 /MISSING LISTWISE /NOTOTAL.
Resources Processor Time	00:00:00.30
Elapsed Time	00:00:01.00

Case Processing Summary

	Cases					
	Valid		Missing		Total	
	N	Percent	N	Percent	N	Percent
F	624	99.7%	2	0.3%	626	100.0%

Descriptives

		Statistic	Std. Error
F	Mean	40.17615%	0.526472%
	95% Confidence Interval for Mean	Lower Bound	39.14228%
		Upper Bound	41.21003%
	5% Trimmed Mean	39.92822%	
	Median	40.28000%	
	Variance	172.956	
	Std. Deviation	13.151269%	

Minimum	5.000%	
Maximum	81.900%	
Range	76.900%	
Interquartile Range		18.010%
Skewness	.201	.098
Kurtosis	-.043	.195

Percentiles

	Percentiles						
	5	10	25	50	75	90	95
Weighted Average(Definition 1)			F	19.00000%	22.22000%	30.60000%	40.28000%
	48.61000%	56.94000%	62.87500%				
Tukey's HingesF			30.60000%	40.28000%	48.61000%		

F

F Stem-and-Leaf Plot

```
 Frequency    Stem &  Leaf

     .00       0 .
    3.00       0 .  568
    4.00       1 .  1223
   32.00       1 .  55555556667888888888999999999
   33.00       2 .  000000111112222222222222233333444
   67.00       2 .
555555555555555666666666666666667777777778888888889999999999999999
   79.00       3 .
0000000000000000001111111111111112222222222223333333333333333333444444444444
44444
   82.00       3 .
555555556666666666666666666666666666677777777777778888888888888888888888888
89999999
   95.00       4 .
00000000000000000000000011111111111111111111112222222333333333333333333333333
33334444444444444444
   86.00       4 .
555555555555555555555666666666666677777777777777777777777788888888888888888
889999999999
   67.00       5 .
000000000000000000001111111111111111111122222222223344444444444444444
   36.00       5 .  55556666666666667778888888888888999
   16.00       6 .  0001112223334444
   12.00       6 .  555688889999
```

```
    9.00       7 .  012223444
    3.00 Extremes    (>=78)

 Stem width:  10.000
 Each leaf:       1 case(s)


DATASET ACTIVATE DataSet1.
DATASET CLOSE DataSet4.
DATASET ACTIVATE DataSet3.
DATASET CLOSE DataSet1.
NEW FILE.
DATASET NAME DataSet5 WINDOW=FRONT.
PRESERVE.
SET DECIMAL DOT.

GET DATA  /TYPE=TXT
  /FILE="/Users/sbateup/Desktop/Loglinear data.csv"
  /ENCODING='UTF8'
  /DELCASE=LINE
  /DELIMITERS=" ,"
  /ARRANGEMENT=DELIMITED
  /FIRSTCASE=2
  /DATATYPEMIN PERCENTAGE=95.0
  /VARIABLES=
  Ref AUTO
  A AUTO
  F AUTO
  Quartile AUTO
  T AUTO
  IAPT AUTO
  Recovery AUTO
  P_A AUTO
  Severity AUTO
  P AUTO
  Gender AUTO
  V12 AUTO
  V13 AUTO
  V14 AUTO
  V15 AUTO
  V16 AUTO
  V17 AUTO
  V18 AUTO
  V19 AUTO
  V20 AUTO
  V21 AUTO
  /MAP.
```

```
RESTORE.

CACHE.
EXECUTE.

Data written to the working file.
21 variables and 627 cases written.
Variable: Ref          Type: String  Format : A12       One or more values were truncated.
Variable: A            Type: String  Format : A10       One or more values were truncated.
Variable: F            Type: String  Format : A10
Variable: Quartile     Type: String  Format : A11
Variable: T            Type: String  Format : A8
Variable: IAPT         Type: String  Format : A6
Variable: Recovery     Type: String  Format : A6
Variable: P_A          Type: Number  Format : F6.3      One or more values were set to
system-missing.
Variable: Severity     Type: Number  Format : F1        One or more values were set to
system-missing.
Variable: P            Type: Number  Format : F1        One or more values were set to system-
missing.
Variable: Gender       Type: Number  Format : F1        One or more values were set to
system-missing.
Variable: V12          Type: Number  Format : F2        One or more values were set to
system-missing.
Variable: V13          Type: Number  Format : F3        One or more values were set to
system-missing.
Variable: V14          Type: Number  Format : F9        One or more values were set to
system-missing.
Variable: V15          Type: Number  Format : F9        One or more values were set to
system-missing.
Variable: V16          Type: Number  Format : F4        One or more values were set to
system-missing.
Variable: V17          Type: Number  Format : F6.3      One or more values were set to
system-missing.
Variable: V18          Type: Number  Format : F6.3      One or more values were set to
system-missing.
Variable: V19          Type: Number  Format : F1        One or more values were set to
system-missing.
Variable: V20          Type: Number  Format : F1        One or more values were set to
system-missing.
Variable: V21          Type: Number  Format : F1        One or more values were set to
system-missing.

DATASET NAME DataSet6 WINDOW=FRONT.
DATASET ACTIVATE DataSet3.
DATASET CLOSE DataSet6.
NEW FILE.
```

```
DATASET NAME DataSet7 WINDOW=FRONT.
PRESERVE.
SET DECIMAL DOT.

GET DATA  /TYPE=TXT
  /FILE="/Users/sbateup/Desktop/LogLinear_clean_240119.csv"
  /ENCODING='UTF8'
  /DELCASE=LINE
  /DELIMITERS=","
  /ARRANGEMENT=DELIMITED
  /FIRSTCASE=1
  /DATATYPEMIN PERCENTAGE=95.0
  /VARIABLES=
  V1 AUTO
  V2 AUTO
  V3 AUTO
  V4 AUTO
  V5 AUTO
  V6 AUTO
  V7 AUTO
  V8 AUTO
  V9 AUTO
  V10 AUTO
  V11 AUTO
  V12 AUTO
  V13 AUTO
  /MAP.

>Warning.  Command name: GET DATA
>(2279) An unexpected/invalid character was encountered in case 131, at or
>around position 9.  This character was skipped.

>Warning.  Command name: GET DATA
>(2279) An unexpected/invalid character was encountered in case 274, at or
>around position 11.  This character was skipped.
RESTORE.

CACHE.
EXECUTE.

>Warning.  Command name: EXECUTE
>(2279) An unexpected/invalid character was encountered in case 131, at or
>around position 9.  This character was skipped.

>Warning.  Command name: EXECUTE
>(2279) An unexpected/invalid character was encountered in case 274, at or
>around position 11.  This character was skipped.
```

Data written to the working file.
13 variables and 610 cases written.
Variable: V1 Type: Number Format : F3 One or more values were set to system-missing.
Variable: V2 Type: String Format : A12 One or more values were truncated.
Variable: V3 Type: Number Format : F1 One or more values were set to system-missing.
Variable: V4 Type: Number Format : F4.2 One or more values were set to system-missing.
Variable: V5 Type: Number Format : F2 One or more values were set to system-missing.
Variable: V6 Type: Number Format : F2 One or more values were set to system-missing.
Variable: V7 Type: Number Format : F2 One or more values were set to system-missing.
Variable: V8 Type: Number Format : F1 One or more values were set to system-missing.
Variable: V9 Type: Number Format : F1 One or more values were set to system-missing.
Variable: V10 Type: Number Format : F1 One or more values were set to system-missing.
Variable: V11 Type: Number Format : F1 One or more values were set to system-missing.
Variable: V12 Type: Number Format : F1 One or more values were set to system-missing.
Variable: V13 Type: Number Format : F1 One or more values were set to system-missing.

```
DATASET NAME DataSet8 WINDOW=FRONT.

SAVE OUTFILE='/Users/sbateup/Desktop/Loglinear.sav'
  /COMPRESSED.
HILOGLINEAR A(0 1) FQuartile(1 4) AgeBinary(0 1) RevoveryD(0 1) RecoveryA(0 1)
GenderX(0 1)
  /METHOD=BACKWARD
  /CRITERIA MAXSTEPS(10) P(.05) ITERATION(20) DELTA(.5)
  /PRINT=FREQ RESID ASSOCIATION ESTIM
  /DESIGN.
```

Hierarchical Loglinear Analysis
Notes
Output Created 25-JAN-2019 16:14:26
Comments
Input Data /Users/sbateup/Desktop/Loglinear.sav
Active Dataset DataSet8
Filter <none>
Weight<none>

Split File <none>
N of Rows in Working Data File 610
Missing Value Handling Definition of Missing User-defined missing values are treated as missing.
Cases Used Statistics are based on all cases with valid data for all variables in the model.
Syntax HILOGLINEAR A(0 1) FQuartile(1 4) AgeBinary(0 1) RevoveryD(0 1) RecoveryA(0 1) GenderX(0 1)
/METHOD=BACKWARD
/CRITERIA MAXSTEPS(10) P(.05) ITERATION(20) DELTA(.5)
/PRINT=FREQ RESID ASSOCIATION ESTIM
/DESIGN.
Resources Processor Time 00:00:00.11
Elapsed Time 00:00:00.00

[DataSet8] /Users/sbateup/Desktop/Loglinear.sav
Warnings
For Design 1, .500 has been added to all observed cells for this saturated model, This value may be changed by using the CRITERIA = DELTA subcommand.

Data Information

		N
Cases	Valid	591
	Out of Rangea	0
	Missing	19
	Weighted Valid	591
Categories	A	2
	FQuartile	4
	AgeBinary	2
	RevoveryD	2
	RecoveryA	2
	GenderX	2

a Cases rejected because of out of range factor values.

Design 1
Convergence Information
Generating Class A*FQuartile*AgeBinary*RevoveryD*RecoveryA*GenderX
Number of Iterations 1
Max. Difference between Observed and Fitted Marginals .000
Convergence Criterion .250

Cell Counts and Residuals

A	FQuartile		AgeBinary		RevoveryD		RecoveryA		GenderX		Observed
		Expected			Residuals		Std. Residuals				
						Counta	%	Count	%		
0	1	0	0	0	0	11.500	1.9%	11.500	1.9%	.000	.000
					1	21.500	3.6%	21.500	3.6%	.000	.000
				1	0	1.500	0.3%	1.500	0.3%	.000	.000
					1	5.500	0.9%	5.500	0.9%	.000	.000
			1	0	0	3.500	0.6%	3.500	0.6%	.000	.000
					1	6.500	1.1%	6.500	1.1%	.000	.000
				1	0	10.500	1.8%	10.500	1.8%	.000	.000
					1	17.500	3.0%	17.500	3.0%	.000	.000
		1	0	0	0	10.500	1.8%	10.500	1.8%	.000	.000
					1	5.500	0.9%	5.500	0.9%	.000	.000
				1	0	4.500	0.8%	4.500	0.8%	.000	.000
					1	12.500	2.1%	12.500	2.1%	.000	.000
			1	0	0	3.500	0.6%	3.500	0.6%	.000	.000
					1	2.500	0.4%	2.500	0.4%	.000	.000
				1	0	4.500	0.8%	4.500	0.8%	.000	.000
					1	17.500	3.0%	17.500	3.0%	.000	.000
	2	0	0	0	0	2.500	0.4%	2.500	0.4%	.000	.000
					1	11.500	1.9%	11.500	1.9%	.000	.000
				1	0	2.500	0.4%	2.500	0.4%	.000	.000
					1	4.500	0.8%	4.500	0.8%	.000	.000
			1	0	0	1.500	0.3%	1.500	0.3%	.000	.000
					1	3.500	0.6%	3.500	0.6%	.000	.000
				1	0	6.500	1.1%	6.500	1.1%	.000	.000
					1	16.500	2.8%	16.500	2.8%	.000	.000
		1	0	0	0	.500	0.1%	.500	0.1%	.000	.000
					1	11.500	1.9%	11.500	1.9%	.000	.000
				1	0	.500	0.1%	.500	0.1%	.000	.000
					1	3.500	0.6%	3.500	0.6%	.000	.000
			1	0	0	.500	0.1%	.500	0.1%	.000	.000
					1	4.500	0.8%	4.500	0.8%	.000	.000
				1	0	4.500	0.8%	4.500	0.8%	.000	.000
					1	18.500	3.1%	18.500	3.1%	.000	.000
	3	0	0	0	0	.500	0.1%	.500	0.1%	.000	.000
					1	7.500	1.3%	7.500	1.3%	.000	.000
				1	0	1.500	0.3%	1.500	0.3%	.000	.000
					1	1.500	0.3%	1.500	0.3%	.000	.000
			1	0	0	.500	0.1%	.500	0.1%	.000	.000
					1	2.500	0.4%	2.500	0.4%	.000	.000
				1	0	3.500	0.6%	3.500	0.6%	.000	.000
					1	5.500	0.9%	5.500	0.9%	.000	.000

		1	0	0	0	1.500	0.3%	1.500	0.3%	.000	.000
					1	1.500	0.3%	1.500	0.3%	.000	.000
				1	0	1.500	0.3%	1.500	0.3%	.000	.000
					1	2.500	0.4%	2.500	0.4%	.000	.000
			1	0	0	1.500	0.3%	1.500	0.3%	.000	.000
					1	1.500	0.3%	1.500	0.3%	.000	.000
				1	0	3.500	0.6%	3.500	0.6%	.000	.000
					1	9.500	1.6%	9.500	1.6%	.000	.000
	4	0	0	0	0	.500	0.1%	.500	0.1%	.000	.000
					1	7.500	1.3%	7.500	1.3%	.000	.000
				1	0	.500	0.1%	.500	0.1%	.000	.000
					1	3.500	0.6%	3.500	0.6%	.000	.000
			1	0	0	.500	0.1%	.500	0.1%	.000	.000
					1	3.500	0.6%	3.500	0.6%	.000	.000
				1	0	1.500	0.3%	1.500	0.3%	.000	.000
					1	6.500	1.1%	6.500	1.1%	.000	.000
		1	0	0	0	1.500	0.3%	1.500	0.3%	.000	.000
					1	4.500	0.8%	4.500	0.8%	.000	.000
				1	0	.500	0.1%	.500	0.1%	.000	.000
					1	2.500	0.4%	2.500	0.4%	.000	.000
			1	0	0	1.500	0.3%	1.500	0.3%	.000	.000
					1	1.500	0.3%	1.500	0.3%	.000	.000
				1	0	2.500	0.4%	2.500	0.4%	.000	.000
					1	5.500	0.9%	5.500	0.9%	.000	.000
1	1	0	0	0	0	.500	0.1%	.500	0.1%	.000	.000
					1	.500	0.1%	.500	0.1%	.000	.000
				1	0	.500	0.1%	.500	0.1%	.000	.000
					1	1.500	0.3%	1.500	0.3%	.000	.000
			1	0	0	1.500	0.3%	1.500	0.3%	.000	.000
					1	.500	0.1%	.500	0.1%	.000	.000
				1	0	.500	0.1%	.500	0.1%	.000	.000
					1	4.500	0.8%	4.500	0.8%	.000	.000
		1	0	0	0	.500	0.1%	.500	0.1%	.000	.000
					1	1.500	0.3%	1.500	0.3%	.000	.000
				1	0	1.500	0.3%	1.500	0.3%	.000	.000
					1	.500	0.1%	.500	0.1%	.000	.000
			1	0	0	.500	0.1%	.500	0.1%	.000	.000
					1	.500	0.1%	.500	0.1%	.000	.000
				1	0	.500	0.1%	.500	0.1%	.000	.000
					1	1.500	0.3%	1.500	0.3%	.000	.000
	2	0	0	0	0	4.500	0.8%	4.500	0.8%	.000	.000
					1	11.500	1.9%	11.500	1.9%	.000	.000
				1	0	3.500	0.6%	3.500	0.6%	.000	.000
					1	2.500	0.4%	2.500	0.4%	.000	.000
			1	0	0	3.500	0.6%	3.500	0.6%	.000	.000
					1	5.500	0.9%	5.500	0.9%	.000	.000
				1	0	3.500	0.6%	3.500	0.6%	.000	.000

				1	6.500	1.1%	6.500	1.1%	.000	.000
	1	0	0	0	1.500	0.3%	1.500	0.3%	.000	.000
				1	8.500	1.4%	8.500	1.4%	.000	.000
			1	0	1.500	0.3%	1.500	0.3%	.000	.000
				1	3.500	0.6%	3.500	0.6%	.000	.000
		1	0	0	1.500	0.3%	1.500	0.3%	.000	.000
				1	.500	0.1%	.500	0.1%	.000	.000
			1	0	1.500	0.3%	1.500	0.3%	.000	.000
				1	9.500	1.6%	9.500	1.6%	.000	.000
3	0	0	0	0	4.500	0.8%	4.500	0.8%	.000	.000
				1	12.500	2.1%	12.500	2.1%	.000	.000
			1	0	1.500	0.3%	1.500	0.3%	.000	.000
				1	5.500	0.9%	5.500	0.9%	.000	.000
		1	0	0	3.500	0.6%	3.500	0.6%	.000	.000
				1	2.500	0.4%	2.500	0.4%	.000	.000
			1	0	4.500	0.8%	4.500	0.8%	.000	.000
				1	19.500	3.3%	19.500	3.3%	.000	.000
	1	0	0	0	3.500	0.6%	3.500	0.6%	.000	.000
				1	8.500	1.4%	8.500	1.4%	.000	.000
			1	0	2.500	0.4%	2.500	0.4%	.000	.000
				1	8.500	1.4%	8.500	1.4%	.000	.000
		1	0	0	4.500	0.8%	4.500	0.8%	.000	.000
				1	5.500	0.9%	5.500	0.9%	.000	.000
			1	0	5.500	0.9%	5.500	0.9%	.000	.000
				1	17.500	3.0%	17.500	3.0%	.000	.000
4	0	0	0	0	5.500	0.9%	5.500	0.9%	.000	.000
				1	16.500	2.8%	16.500	2.8%	.000	.000
			1	0	3.500	0.6%	3.500	0.6%	.000	.000
				1	7.500	1.3%	7.500	1.3%	.000	.000
		1	0	0	5.500	0.9%	5.500	0.9%	.000	.000
				1	7.500	1.3%	7.500	1.3%	.000	.000
			1	0	7.500	1.3%	7.500	1.3%	.000	.000
				1	15.500	2.6%	15.500	2.6%	.000	.000
	1	0	0	0	7.500	1.3%	7.500	1.3%	.000	.000
				1	11.500	1.9%	11.500	1.9%	.000	.000
			1	0	1.500	0.3%	1.500	0.3%	.000	.000
				1	7.500	1.3%	7.500	1.3%	.000	.000
		1	0	0	2.500	0.4%	2.500	0.4%	.000	.000
				1	9.500	1.6%	9.500	1.6%	.000	.000
			1	0	3.500	0.6%	3.500	0.6%	.000	.000
				1	24.500	4.1%	24.500	4.1%	.000	.000

a For saturated models, .500 has been added to all observed cells.

Goodness-of-Fit Tests

	Chi-Square	df	Sig.
Likelihood Ratio	.000	0	.
Pearson	.000	0	.

K-Way and Higher-Order Effects

K	df	Likelihood Ratio		Pearson		Number of Iterations
		Chi-Square	Sig.	Chi-Square	Sig.	
K-way and Higher Order Effectsa	1	127	625.044	.000	699.613	.000
0						
2	119	465.134	.000	485.295	.000	2
3	94	101.549	.279	94.161	.476	8
4	54	52.617	.528	49.863	.635	8
5	19	20.084	.390	16.298	.637	8
6	3	.438	.932	.230	.973	4
K-way Effectsb	1	8	159.910	.000	214.318	.000
0						
2	25	363.585	.000	391.134	.000	0
3	40	48.932	.157	44.298	.295	0
4	35	32.533	.588	33.565	.537	0
5	16	19.645	.237	16.068	.448	0
6	3	.438	.932	.230	.973	0

df used for these tests have NOT been adjusted for structural or sampling zeros. Tests using these df may be conservative.

a Tests that k-way and higher order effects are zero.

b Tests that k-way effects are zero.

Partial Associations

Effect	df	Partial Chi-Square	Sig.	Number of Iterations
A*FQuartile*AgeBinary*RevoveryD*RecoveryA	3	2.058	.560	5
A*FQuartile*AgeBinary*RevoveryD*GenderX	3	.651	.885	4
A*FQuartile*AgeBinary*RecoveryA*GenderX	3	7.437	.059	7
A*FQuartile*RevoveryD*RecoveryA*GenderX	3	.199	.978	4
A*AgeBinary*RevoveryD*RecoveryA*GenderX	1	.028	.867	4
FQuartile*AgeBinary*RevoveryD*RecoveryA*GenderX	3	2.392	.495	3
A*FQuartile*AgeBinary*RevoveryD	3	1.988	.575	9
A*FQuartile*AgeBinary*RecoveryA	3	5.201	.158	8
A*FQuartile*RevoveryD*RecoveryA	3	.378	.945	8
A*AgeBinary*RevoveryD*RecoveryA	1	.154	.695	8
FQuartile*AgeBinary*RevoveryD*RecoveryA	3	6.482	.090	7
A*FQuartile*AgeBinary*GenderX	3	4.516	.211	8
A*FQuartile*RevoveryD*GenderX	3	3.635	.304	7
A*AgeBinary*RevoveryD*GenderX	1	.758	.384	6

FQuartile*AgeBinary*RevoveryD*GenderX	3	4.751	.191	6
A*FQuartile*RecoveryA*GenderX	3	1.687	.640	7
A*AgeBinary*RecoveryA*GenderX	1	3.435	.064	5
FQuartile*AgeBinary*RecoveryA*GenderX	3	1.834	.607	5
A*RevoveryD*RecoveryA*GenderX	1	3.522	.061	7
FQuartile*RevoveryD*RecoveryA*GenderX	3	2.626	.453	8
AgeBinary*RevoveryD*RecoveryA*GenderX	1	.003	.955	9
A*FQuartile*AgeBinary	3	1.177	.759	8
A*FQuartile*RevoveryD	3	1.217	.749	8
A*AgeBinary*RevoveryD	1	1.000	.317	8
FQuartile*AgeBinary*RevoveryD	3	4.486	.214	7
A*FQuartile*RecoveryA	3	2.868	.412	8
A*AgeBinary*RecoveryA	1	.131	.718	7
FQuartile*AgeBinary*RecoveryA	3	4.016	.260	7
A*RevoveryD*RecoveryA	1	1.340	.247	8
FQuartile*RevoveryD*RecoveryA	3	.604	.896	8
AgeBinary*RevoveryD*RecoveryA	1	.463	.496	9
A*FQuartile*GenderX	3	1.913	.591	8
A*AgeBinary*GenderX	1	1.594	.207	8
FQuartile*AgeBinary*GenderX	3	6.719	.081	8
A*RevoveryD*GenderX	1	.505	.477	8
FQuartile*RevoveryD*GenderX	3	.767	.857	8
AgeBinary*RevoveryD*GenderX	1	.395	.530	8
A*RecoveryA*GenderX	1	3.191	.074	8
FQuartile*RecoveryA*GenderX	3	7.580	.056	8
AgeBinary*RecoveryA*GenderX	1	3.463	.063	8
RevoveryD*RecoveryA*GenderX	1	1.902	.168	8
A*FQuartile	3	218.764	.000	5
A*AgeBinary	1	.053	.818	8
FQuartile*AgeBinary	3	1.912	.591	8
A*RevoveryD	1	.234	.628	6
FQuartile*RevoveryD	3	3.198	.362	6
AgeBinary*RevoveryD	1	.033	.855	8
A*RecoveryA	1	.515	.473	7
FQuartile*RecoveryA	3	2.238	.525	6
AgeBinary*RecoveryA	1	5.302	.021	8
RevoveryD*RecoveryA	1	116.840	.000	6
A*GenderX	1	1.579	.209	8
FQuartile*GenderX	3	8.695	.034	7
AgeBinary*GenderX	1	.380	.538	8
RevoveryD*GenderX	1	1.088	.297	8
RecoveryA*GenderX	1	3.569	.059	7
A	1	.205	.651	2
FQuartile	3	2.788	.425	2
AgeBinary	1	2.847	.092	2
RevoveryD	1	4.759	.029	2
RecoveryA	1	8.074	.004	2

GenderX 1 141.236 .000 2

Parameter Estimates

Effect	Parameter	Estimate	Std. Error	Z	Sig.	95% Confidence Interval			
						Lower Bound	Upper Bound		
A*FQuartile*AgeBinary*RevoveryD*RecoveryA*GenderX	1	.008	.132	.058	.954	-.251	.267		
	2	-.077	.111	-.695	.487	-.295	.141		
	3	.012	.107	.110	.912	-.197	.221		
A*FQuartile*AgeBinary*RevoveryD*RecoveryA	1	.105	.132	.797	.425	-.154	.364		
	2	.002	.111	.014	.989	-.216	.220		
	3	-.077	.107	-.722	.471	-.286	.132		
A*FQuartile*AgeBinary*RevoveryD*GenderX	1	-.015	.132	-.113	.910	-.274	.244		
	2	-.041	.111	-.366	.714	-.259	.177		
	3	-.026	.107	-.244	.807	-.235	.183		
A*FQuartile*AgeBinary*RecoveryA*GenderX	1	-.173	.132	-1.310	.190	-.432	.086		
	2	.220	.111	1.977	.048	.002	.438		
	3	-.107	.107	-1.001	.317	-.315	.102		
A*FQuartile*RevoveryD*RecoveryA*GenderX	1	.119	.132	.897	.370	-.141	.378		
	2	.043	.111	.391	.696	-.175	.261		
	3	-.073	.107	-.689	.491	-.282	.135		
A*AgeBinary*RevoveryD*RecoveryA*GenderX	1	-.017	.067	-.250	.802	-.147	.114		
FQuartile*AgeBinary*RevoveryD*RecoveryA*GenderX	1	.057	.132	.428	.669	-.202	.316		
	2	.052	.111	.471	.637	-.166	.270		
	3	-.064	.107	-.606	.545	-.273	.144		
A*FQuartile*AgeBinary*RevoveryD	1	-.013	.132	-.101	.920	-.272	.246		
	2	.085	.111	.764	.445	-.133	.303		
	3	-.036	.107	-.335	.738	-.244	.173		
A*FQuartile*AgeBinary*RecoveryA	1	.133	.132	1.005	.315	-.126	.392		
	2	-.114	.111	-1.027	.304	-.332	.104		
	3	.000	.107	.004	.996	-.208	.209		
A*FQuartile*RevoveryD*RecoveryA	1	.133	.132	1.003	.316	-.126	.392		
	2	.019	.111	.171	.864	-.199	.237		
	3	-.089	.107	-.840	.401	-.298	.119		
A*AgeBinary*RevoveryD*RecoveryA	1	.031	.067	.465	.642	-.100	.162		
FQuartile*AgeBinary*RevoveryD*RecoveryA	1	.060	.132	.455	.649	-.199	.319		
	2	-.093	.111	-.836	.403	-.311	.125		

3 .105 .107 .983 .325 -.104 .313
A*FQuartile*AgeBinary*GenderX 1 .043 .132 .322 .747 -.216 .302
2 .158 .111 1.423 .155 -.060 .376
3 -.042 .107 -.398 .690 -.251 .166
A*FQuartile*RevoveryD*GenderX 1 -.012 .132 -.089 .929 -.271 .247
2 .013 .111 .121 .904 -.204 .231
3 .083 .107 .779 .436 -.126 .292
A*AgeBinary*RevoveryD*GenderX 1 .004 .067 .063 .950 -.126 .135
FQuartile*AgeBinary*RevoveryD*GenderX 1 -.063 .132 -.479 .632 -.322
.196
2 .124 .111 1.112 .266 -.094 .342
3 -.033 .107 -.310 .757 -.242 .176
A*FQuartile*RecoveryA*GenderX 1 .103 .132 .781 .435 -.156 .362
2 -.035 .111 -.310 .756 -.252 .183
3 -.096 .107 -.897 .370 -.304 .113
A*AgeBinary*RecoveryA*GenderX 1 -.086 .067 -1.289 .197 -.217 .045
FQuartile*AgeBinary*RecoveryA*GenderX 1 .148 .132 1.122 .262 -.111
.407
2 7.466E-5 .111 .001 .999 -.218 .218
3 -.079 .107 -.745 .456 -.288 .129
A*RevoveryD*RecoveryA*GenderX 1 .077 .067 1.151 .250 -.054 .207
FQuartile*RevoveryD*RecoveryA*GenderX 1 -.020 .132 -.153 .879 -.279
.239
2 -.082 .111 -.734 .463 -.300 .136
3 -.003 .107 -.027 .978 -.212 .206
AgeBinary*RevoveryD*RecoveryA*GenderX 1 .003 .067 .045 .964 -.128
.134
A*FQuartile*AgeBinary 1 .032 .132 .240 .810 -.227 .291
2 -.009 .111 -.078 .938 -.227 .209
3 .034 .107 .318 .750 -.175 .243
A*FQuartile*RevoveryD 1 .104 .132 .789 .430 -.155 .363
2 -.123 .111 -1.107 .268 -.341 .095
3 .000 .107 .002 .998 -.209 .209
A*AgeBinary*RevoveryD 1 .045 .067 .681 .496 -.085 .176
FQuartile*AgeBinary*RevoveryD 1 -.153 .132 -1.159 .247 -.412 .106
2 .038 .111 .346 .729 -.179 .256
3 .082 .107 .769 .442 -.127 .291
A*FQuartile*RecoveryA 1 .136 .132 1.030 .303 -.123 .395
2 -.067 .111 -.600 .549 -.285 .151
3 -.066 .107 -.617 .537 -.274 .143
A*AgeBinary*RecoveryA 1 .010 .067 .154 .877 -.120 .141
FQuartile*AgeBinary*RecoveryA 1 .023 .132 .172 .864 -.236 .282
2 .074 .111 .661 .508 -.144 .291
3 -.001 .107 -.012 .990 -.210 .207
A*RevoveryD*RecoveryA 1 .074 .067 1.111 .266 -.057 .205
FQuartile*RevoveryD*RecoveryA 1 -.086 .132 -.651 .515 -.345 .173
2 .093 .111 .834 .405 -.125 .311

	3	-.007	.107	-.064	.949	-.216	.202			
AgeBinary*RevoveryD*RecoveryA	1	.008	.067	.113	.910	-.123	.138			
A*FQuartile*GenderX	1	.053	.132	.403	.687	-.206	.312			
	2	-.148	.111	-1.327	.185	-.365	.070			
	3	.109	.107	1.028	.304	-.099	.318			
A*AgeBinary*GenderX	1	-.060	.067	-.901	.368	-.191	.071			
FQuartile*AgeBinary*GenderX	1	-.073	.132	-.552	.581	-.332	.186			
	2	.215	.111	1.933	.053	-.003	.433			
	3	-.088	.107	-.827	.408	-.297	.121			
A*RevoveryD*GenderX	1	-.027	.067	-.408	.683	-.158	.103			
FQuartile*RevoveryD*GenderX	1	.085	.132	.642	.521	-.174	.344			
	2	-.044	.111	-.395	.693	-.262	.174			
	3	-.011	.107	-.099	.921	-.219	.198			
AgeBinary*RevoveryD*GenderX	1	-.016	.067	-.242	.809	-.147	.115			
A*RecoveryA*GenderX	1	-.084	.067	-1.253	.210	-.214	.047			
FQuartile*RecoveryA*GenderX	1	.161	.132	1.218	.223	-.098	.420			
	2	-.123	.111	-1.110	.267	-.341	.094			
	3	-.033	.107	-.310	.757	-.242	.176			
AgeBinary*RecoveryA*GenderX	1	-.064	.067	-.960	.337	-.195	.067			
RevoveryD*RecoveryA*GenderX	1	-.128	.067	-1.918	.055	-.258	.003			
A*FQuartile	1	1.069	.132	8.085	.000	.809	1.328			
	2	.034	.111	.308	.758	-.184	.252			
	3	-.470	.107	-4.412	.000	-.679	-.261			
A*AgeBinary	1	-.024	.067	-.364	.716	-.155	.106			
FQuartile*AgeBinary	1	-.010	.132	-.072	.943	-.269	.250			
	2	.233	.111	2.095	.036	.015	.451			
	3	-.166	.107	-1.560	.119	-.375	.043			
A*RevoveryD	1	-.044	.067	-.660	.509	-.175	.087			
FQuartile*RevoveryD	1	.075	.132	.566	.572	-.184	.334			
	2	.033	.111	.294	.768	-.185	.251			
	3	-.075	.107	-.705	.481	-.284	.134			
AgeBinary*RevoveryD	1	-.021	.067	-.311	.756	-.151	.110			
A*RecoveryA	1	-.060	.067	-.899	.368	-.191	.071			
FQuartile*RecoveryA	1	-.004	.132	-.033	.973	-.263	.255			
	2	-.029	.111	-.262	.793	-.247	.189			
	3	-.088	.107	-.827	.408	-.297	.121			
AgeBinary*RecoveryA	1	.052	.067	.777	.437	-.079	.182			
RevoveryD*RecoveryA	1	.361	.067	5.423	.000	.231	.492			
A*GenderX	1	-.089	.067	-1.342	.180	-.220	.041			
FQuartile*GenderX	1	.213	.132	1.615	.106	-.046	.472			
	2	-.094	.111	-.846	.398	-.312	.124			
	3	.043	.107	.405	.686	-.166	.252			
AgeBinary*GenderX	1	-.013	.067	-.198	.843	-.144	.117			
RevoveryD*GenderX	1	-.055	.067	-.828	.408	-.186	.075			
RecoveryA*GenderX	1	.065	.067	.972	.331	-.066	.195			
A	1	-.007	.067	-.099	.921	-.137	.124			
FQuartile	1	-.287	.132	-2.169	.030	-.546	-.028			

	2	.065	.111	.583	.560	-.153	.283
	3	.076	.107	.710	.478	-.133	.284
AgeBinary	1	.086	.067	1.285	.199	-.045	.216
RevoveryD	1	-.083	.067	-1.246	.213	-.214	.048
RecoveryA	1	-.125	.067	-1.882	.060	-.256	.005
GenderX	1	-.456	.067	-6.835	.000	-.586	-.325

Backward Elimination Statistics

Step Summary

Stepa		Effects	Chi-Squarec	df	Sig.	Number of Iterations
0	Generating Classb					
	A*FQuartile*AgeBinary*RevoveryD*RecoveryA*GenderX		.000	0	.	
	Deleted Effect 1	A*FQuartile*AgeBinary*RevoveryD*RecoveryA*GenderX	.438	3	.932	4
1	Generating Classb	A*FQuartile*AgeBinary*RevoveryD*RecoveryA, A*FQuartile*AgeBinary*RevoveryD*GenderX, A*FQuartile*AgeBinary*RecoveryA*GenderX, A*FQuartile*RevoveryD*RecoveryA*GenderX, A*AgeBinary*RevoveryD*RecoveryA*GenderX, FQuartile*AgeBinary*RevoveryD*RecoveryA*GenderX	.438	3	.932	
	Deleted Effect 1	A*FQuartile*AgeBinary*RevoveryD*RecoveryA	2.058	3	.560	5
	2	A*FQuartile*AgeBinary*RevoveryD*GenderX	.651	3	.885	4
	3	A*FQuartile*AgeBinary*RecoveryA*GenderX	7.437	3	.059	7
	4	A*FQuartile*RevoveryD*RecoveryA*GenderX	.199	3	.978	4
	5	A*AgeBinary*RevoveryD*RecoveryA*GenderX	.028	1	.867	4
	6	FQuartile*AgeBinary*RevoveryD*RecoveryA*GenderX	2.392	3	.495	3
2	Generating Classb	A*FQuartile*AgeBinary*RevoveryD*RecoveryA, A*FQuartile*AgeBinary*RevoveryD*GenderX, A*FQuartile*AgeBinary*RecoveryA*GenderX, A*AgeBinary*RevoveryD*RecoveryA*GenderX, FQuartile*AgeBinary*RevoveryD*RecoveryA*GenderX	.637	6	.996	
	Deleted Effect 1	A*FQuartile*AgeBinary*RevoveryD*RecoveryA	2.435	3	.487	5
	2	A*FQuartile*AgeBinary*RevoveryD*GenderX	.881	3	.830	5
	3	A*FQuartile*AgeBinary*RecoveryA*GenderX	12.842	3	.005	8
	4	A*AgeBinary*RevoveryD*RecoveryA*GenderX	.034	1	.854	4
	5	FQuartile*AgeBinary*RevoveryD*RecoveryA*GenderX	2.320	3	.509	4

3 Generating Classb A*FQuartile*AgeBinary*RevoveryD*RecoveryA, A*FQuartile*AgeBinary*RevoveryD*GenderX, A*FQuartile*AgeBinary*RecoveryA*GenderX, FQuartile*AgeBinary*RevoveryD*RecoveryA*GenderX, A*RevoveryD*RecoveryA*GenderX .671 7 .999

Deleted Effect 1 A*FQuartile*AgeBinary*RevoveryD*RecoveryA 2.336 3 .506 6

2 A*FQuartile*AgeBinary*RevoveryD*GenderX 1.095 3 .778 4

3 A*FQuartile*AgeBinary*RecoveryA*GenderX 13.400 3 .004 8

4 FQuartile*AgeBinary*RevoveryD*RecoveryA*GenderX 2.791 3 .425 4

5 A*RevoveryD*RecoveryA*GenderX .129 1 .719 4

4 Generating Classb A*FQuartile*AgeBinary*RevoveryD*RecoveryA, A*FQuartile*AgeBinary*RecoveryA*GenderX, FQuartile*AgeBinary*RevoveryD*RecoveryA*GenderX, A*RevoveryD*RecoveryA*GenderX, A*FQuartile*RevoveryD*GenderX, A*AgeBinary*RevoveryD*GenderX 1.766 10 .998

Deleted Effect 1 A*FQuartile*AgeBinary*RevoveryD*RecoveryA 3.846 3 .279 6

2 A*FQuartile*AgeBinary*RecoveryA*GenderX 14.987 3 .002 7

3 FQuartile*AgeBinary*RevoveryD*RecoveryA*GenderX 2.210 3 .530 5

4 A*RevoveryD*RecoveryA*GenderX .214 1 .644 5

5 A*FQuartile*RevoveryD*GenderX 4.135 3 .247 5

6 A*AgeBinary*RevoveryD*GenderX .299 1 .584 4

5 Generating Classb A*FQuartile*AgeBinary*RevoveryD*RecoveryA, A*FQuartile*AgeBinary*RecoveryA*GenderX, FQuartile*AgeBinary*RevoveryD*RecoveryA*GenderX, A*FQuartile*RevoveryD*GenderX, A*AgeBinary*RevoveryD*GenderX 1.980 11 .999

Deleted Effect 1 A*FQuartile*AgeBinary*RevoveryD*RecoveryA 5.405 3 .144 6

2 A*FQuartile*AgeBinary*RecoveryA*GenderX 17.715 3 .001 7

3 FQuartile*AgeBinary*RevoveryD*RecoveryA*GenderX 2.858 3 .414 5

4 A*FQuartile*RevoveryD*GenderX 3.958 3 .266 6

5 A*AgeBinary*RevoveryD*GenderX .262 1 .609 5

6 Generating Classb A*FQuartile*AgeBinary*RevoveryD*RecoveryA, A*FQuartile*AgeBinary*RecoveryA*GenderX, FQuartile*AgeBinary*RevoveryD*RecoveryA*GenderX, A*FQuartile*RevoveryD*GenderX 2.242 12 .999

Deleted Effect 1 A*FQuartile*AgeBinary*RevoveryD*RecoveryA 5.931 3 .115 4

2 A*FQuartile*AgeBinary*RecoveryA*GenderX 17.842 3 .000 6

3 FQuartile*AgeBinary*RevoveryD*RecoveryA*GenderX 2.766 3
.429 5
4 A*FQuartile*RevoveryD*GenderX 4.007 3 .261 8
7 Generating Classb A*FQuartile*AgeBinary*RevoveryD*RecoveryA, A*FQuartile*AgeBinary*RecoveryA*GenderX, A*FQuartile*RevoveryD*GenderX, FQuartile*AgeBinary*RevoveryD*GenderX, FQuartile*RevoveryD*RecoveryA*GenderX, AgeBinary*RevoveryD*RecoveryA*GenderX 5.007 15 .992
Deleted Effect 1 A*FQuartile*AgeBinary*RevoveryD*RecoveryA 5.130 3
.163 4
2 A*FQuartile*AgeBinary*RecoveryA*GenderX 16.492 3
.001 6
3 A*FQuartile*RevoveryD*GenderX 4.366 3 .225 7
4 FQuartile*AgeBinary*RevoveryD*GenderX 3.683 3 .298 4
5 FQuartile*RevoveryD*RecoveryA*GenderX 3.249 3 .355 4
6 AgeBinary*RevoveryD*RecoveryA*GenderX .173 1 .677 4
8 Generating Classb A*FQuartile*AgeBinary*RevoveryD*RecoveryA, A*FQuartile*AgeBinary*RecoveryA*GenderX, A*FQuartile*RevoveryD*GenderX, FQuartile*AgeBinary*RevoveryD*GenderX, FQuartile*RevoveryD*RecoveryA*GenderX
5.181 16 .995
Deleted Effect 1 A*FQuartile*AgeBinary*RevoveryD*RecoveryA 4.960 3
.175 4
2 A*FQuartile*AgeBinary*RecoveryA*GenderX 16.500 3
.001 6
3 A*FQuartile*RevoveryD*GenderX 4.203 3 .240 7
4 FQuartile*AgeBinary*RevoveryD*GenderX 3.733 3 .292 4
5 FQuartile*RevoveryD*RecoveryA*GenderX 3.210 3 .360 4
9 Generating Classb A*FQuartile*AgeBinary*RevoveryD*RecoveryA, A*FQuartile*AgeBinary*RecoveryA*GenderX, A*FQuartile*RevoveryD*GenderX, FQuartile*AgeBinary*RevoveryD*GenderX, RevoveryD*RecoveryA*GenderX 8.391 19
.982
Deleted Effect 1 A*FQuartile*AgeBinary*RevoveryD*RecoveryA 4.502 3
.212 4
2 A*FQuartile*AgeBinary*RecoveryA*GenderX 15.195 3
.002 6
3 A*FQuartile*RevoveryD*GenderX 4.501 3 .212 7
4 FQuartile*AgeBinary*RevoveryD*GenderX 2.859 3 .414 4
5 RevoveryD*RecoveryA*GenderX .936 1 .333 4
10 Generating Classb A*FQuartile*AgeBinary*RevoveryD*RecoveryA, A*FQuartile*AgeBinary*RecoveryA*GenderX, A*FQuartile*RevoveryD*GenderX, RevoveryD*RecoveryA*GenderX, AgeBinary*RevoveryD*GenderX 11.250 22 .971
a At each step, the effect with the largest significance level for the Likelihood Ratio Change is deleted, provided the significance level is larger than .050.

b Statistics are displayed for the best model at each step after step 0.

c For 'Deleted Effect', this is the change in the Chi-Square after the effect is deleted from the model.

Convergence Informationa
Generating Class A*FQuartile*AgeBinary*RevoveryD*RecoveryA, A*FQuartile*AgeBinary*RecoveryA*GenderX, A*FQuartile*RevoveryD*GenderX, RevoveryD*RecoveryA*GenderX, AgeBinary*RevoveryD*GenderX
Number of Iterations 0
Max. Difference between Observed and Fitted Marginals .000
Convergence Criterion .250
a Statistics for the final model after Backward Elimination.

Cell Counts and Residuals

A	FQuartile	AgeBinary	RevoveryD	RecoveryA	GenderX	Observed	Expected		Residuals		Std. Residuals
							Count	%	Count	%	
0	1	0	0	0	0	11.000	1.9%	10.376	1.8%	.624	.194
					1	21.000	3.6%	21.682	3.7%	-.682	-.146
				1	0	1.000	0.2%	1.624	0.3%	-.624	-.490
					1	5.000	0.8%	4.392	0.7%	.608	.290
			1	0	0	3.000	0.5%	3.697	0.6%	-.697	-.363
					1	6.000	1.0%	5.312	0.9%	.688	.298
				1	0	10.000	1.7%	9.446	1.6%	.554	.180
					1	17.000	2.9%	17.564	3.0%	-.564	-.135
		1	0	0	0	10.000	1.7%	9.995	1.7%	.005	.002
					1	5.000	0.8%	4.935	0.8%	.065	.029
				1	0	4.000	0.7%	4.011	0.7%	-.011	-.005
					1	12.000	2.0%	11.991	2.0%	.009	.003
			1	0	0	3.000	0.5%	2.911	0.5%	.089	.052
					1	2.000	0.3%	2.078	0.4%	-.078	-.054
				1	0	4.000	0.7%	3.944	0.7%	.056	.028
					1	17.000	2.9%	17.050	2.9%	-.050	-.012
	2	0	0	0	0	2.000	0.3%	2.236	0.4%	-.236	-.158
					1	11.000	1.9%	10.742	1.8%	.258	.079
				1	0	2.000	0.3%	1.626	0.3%	.374	.293
					1	4.000	0.7%	4.356	0.7%	-.356	-.171
			1	0	0	1.000	0.2%	.740	0.1%	.260	.302
					1	3.000	0.5%	3.262	0.6%	-.262	-.145
				1	0	6.000	1.0%	6.365	1.1%	-.365	-.145
					1	16.000	2.7%	15.648	2.6%	.352	.089
		1	0	0	0	.000	0.0%	.000	0.0%	.000	.000
					1	11.000	1.9%	11.058	1.9%	-.058	-.017
				1	0	.000	0.0%	.167	0.0%	-.167	-.409
					1	3.000	0.5%	2.845	0.5%	.155	.092

			1	0	0	.000	0.0%	.000	0.0%	.000	.000
					1	4.000	0.7%	3.985	0.7%	.015	.008
				1	0	4.000	0.7%	3.904	0.7%	.096	.049
					1	18.000	3.0%	18.100	3.1%	-.100	-.024
	3	0	0	0	0	.000	0.0%	.000	0.0%	.000	.000
					1	7.000	1.2%	7.013	1.2%	-.013	-.005
				1	0	1.000	0.2%	.863	0.1%	.137	.148
					1	1.000	0.2%	1.133	0.2%	-.133	-.125
			1	0	0	.000	0.0%	.000	0.0%	.000	.000
					1	2.000	0.3%	1.997	0.3%	.003	.002
				1	0	3.000	0.5%	3.146	0.5%	-.146	-.082
					1	5.000	0.8%	4.861	0.8%	.139	.063
		1	0	0	0	1.000	0.2%	1.116	0.2%	-.116	-.109
					1	1.000	0.2%	.886	0.1%	.114	.121
				1	0	1.000	0.2%	1.027	0.2%	-.027	-.027
					1	2.000	0.3%	1.974	0.3%	.026	.019
			1	0	0	1.000	0.2%	.886	0.1%	.114	.121
					1	1.000	0.2%	1.114	0.2%	-.114	-.108
				1	0	3.000	0.5%	2.969	0.5%	.031	.018
					1	9.000	1.5%	9.031	1.5%	-.031	-.010
	4	0	0	0	0	.000	0.0%	.000	0.0%	.000	.000
					1	7.000	1.2%	7.057	1.2%	-.057	-.021
				1	0	.000	0.0%	.029	0.0%	-.029	-.171
					1	3.000	0.5%	3.013	0.5%	-.013	-.007
			1	0	0	.000	0.0%	.000	0.0%	.000	.000
					1	3.000	0.5%	2.974	0.5%	.026	.015
				1	0	1.000	0.2%	1.001	0.2%	-.001	-.001
					1	6.000	1.0%	5.960	1.0%	.040	.016
		1	0	0	0	1.000	0.2%	.850	0.1%	.150	.163
					1	4.000	0.7%	4.051	0.7%	-.051	-.025
				1	0	.000	0.0%	.131	0.0%	-.131	-.362
					1	2.000	0.3%	1.879	0.3%	.121	.088
			1	0	0	1.000	0.2%	1.085	0.2%	-.085	-.081
					1	1.000	0.2%	.974	0.2%	.026	.026
				1	0	2.000	0.3%	1.890	0.3%	.110	.080
					1	5.000	0.8%	5.092	0.9%	-.092	-.041
1	1	0	0	0	0	.000	0.0%	.000	0.0%	.000	.000
					1	.000	0.0%	.000	0.0%	.000	.000
				1	0	.000	0.0%	.000	0.0%	.000	.000
					1	1.000	0.2%	.926	0.2%	.074	.076
			1	0	0	1.000	0.2%	.857	0.1%	.143	.154
					1	.000	0.0%	.000	0.0%	.000	.000
				1	0	.000	0.0%	.000	0.0%	.000	.000
					1	4.000	0.7%	4.124	0.7%	-.124	-.061
		1	0	0	0	.000	0.0%	.000	0.0%	.000	.000
					1	1.000	0.2%	.928	0.2%	.072	.074
				1	0	1.000	0.2%	.994	0.2%	.006	.006

				1	.000	0.0%	.146	0.0%	-.146	-.382
		1	0	0	.000	0.0%	.000	0.0%	.000	.000
				1	.000	0.0%	.000	0.0%	.000	.000
			1	0	.000	0.0%	.145	0.0%	-.145	-.380
				1	1.000	0.2%	.872	0.1%	.128	.137
2	0	0	0	0	4.000	0.7%	4.819	0.8%	-.819	-.373
				1	11.000	1.9%	10.206	1.7%	.794	.249
			1	0	3.000	0.5%	2.319	0.4%	.681	.448
				1	2.000	0.3%	2.696	0.5%	-.696	-.424
		1	0	0	3.000	0.5%	2.209	0.4%	.791	.532
				1	5.000	0.8%	5.787	1.0%	-.787	-.327
			1	0	3.000	0.5%	3.686	0.6%	-.686	-.357
				1	6.000	1.0%	5.303	0.9%	.697	.303
	1	0	0	0	1.000	0.2%	1.547	0.3%	-.547	-.440
				1	8.000	1.4%	7.380	1.2%	.620	.228
			1	0	1.000	0.2%	.285	0.0%	.715	1.338
				1	3.000	0.5%	3.717	0.6%	-.717	-.372
		1	0	0	1.000	0.2%	.369	0.1%	.631	1.037
				1	.000	0.0%	.647	0.1%	-.647	-.805
			1	0	1.000	0.2%	1.727	0.3%	-.727	-.553
				1	9.000	1.5%	8.268	1.4%	.732	.255
3	0	0	0	0	4.000	0.7%	4.564	0.8%	-.564	-.264
				1	12.000	2.0%	11.407	1.9%	.593	.176
			1	0	1.000	0.2%	.573	0.1%	.427	.564
				1	5.000	0.8%	5.447	0.9%	-.447	-.192
		1	0	0	3.000	0.5%	2.401	0.4%	.599	.387
				1	2.000	0.3%	2.625	0.4%	-.625	-.386
			1	0	4.000	0.7%	4.454	0.8%	-.454	-.215
				1	19.000	3.2%	18.518	3.1%	.482	.112
	1	0	0	0	3.000	0.5%	3.251	0.6%	-.251	-.139
				1	8.000	1.4%	7.745	1.3%	.255	.092
			1	0	2.000	0.3%	1.606	0.3%	.394	.311
				1	8.000	1.4%	8.395	1.4%	-.395	-.136
		1	0	0	4.000	0.7%	3.748	0.6%	.252	.130
				1	5.000	0.8%	5.254	0.9%	-.254	-.111
			1	0	5.000	0.8%	5.397	0.9%	-.397	-.171
				1	17.000	2.9%	16.601	2.8%	.399	.098
4	0	0	0	0	5.000	0.8%	5.536	0.9%	-.536	-.228
				1	16.000	2.7%	15.393	2.6%	.607	.155
			1	0	3.000	0.5%	2.434	0.4%	.566	.362
				1	7.000	1.2%	7.537	1.3%	-.537	-.196
		1	0	0	5.000	0.8%	4.428	0.7%	.572	.272
				1	7.000	1.2%	7.599	1.3%	-.599	-.217
			1	0	7.000	1.2%	7.570	1.3%	-.570	-.207
				1	15.000	2.5%	14.466	2.4%	.534	.140
	1	0	0	0	7.000	1.2%	6.611	1.1%	.389	.151
				1	11.000	1.9%	11.458	1.9%	-.458	-.135

	1	0	1.000	0.2%	1.408	0.2%	-.408	-.344
		1	7.000	1.2%	6.612	1.1%	.388	.151
1	0	0	2.000	0.3%	2.446	0.4%	-.446	-.285
		1	9.000	1.5%	8.539	1.4%	.461	.158
	1	0	3.000	0.5%	2.579	0.4%	.421	.262
		1	24.000	4.1%	24.395	4.1%	-.395	-.080

Goodness-of-Fit Tests

	Chi-Square	df	Sig.
Likelihood Ratio	11.250	22	.971
Pearson	11.087	22	.973

```
GENLOG A FQuartile GenderX AgeBinary RevoveryD RecoveryA
  /MODEL=MULTINOMIAL
  /PRINT=FREQ RESID ADJRESID ZRESID DEV DESIGN ITERATION
  /PLOT=RESID(ADJRESID) NORMPROB(ADJRESID)
  /CRITERIA=CIN(95) ITERATE(20) CONVERGE(0.001) DELTA(.5)
  /DESIGN A FQuartile GenderX AgeBinary RevoveryD RecoveryA.
```

General Loglinear

Notes

Output Created		02-FEB-2019 08:30:11
Comments		
Input	Data	/Users/sbateup/Desktop/Loglinear.sav
	Active Dataset	DataSet8
	Filter	<none>
	Weight	<none>
	Split File	<none>
	N of Rows in Working Data File	610
Missing Value Handling	Definition of Missing	User-defined missing values are treated as missing.
	Cases Used	Statistics are based on all cases with valid data for all variables in the model.
Syntax		GENLOG A FQuartile GenderX AgeBinary RevoveryD RecoveryA /MODEL=MULTINOMIAL /PRINT=FREQ RESID ADJRESID ZRESID DEV DESIGN ITERATION /PLOT=RESID(ADJRESID) NORMPROB(ADJRESID) /CRITERIA=CIN(95) ITERATE(20) CONVERGE(0.001) DELTA(.5) /DESIGN A FQuartile GenderX AgeBinary RevoveryD RecoveryA.
Resources	Processor Time	00:00:00.69
	Elapsed Time	00:00:00.00

Data Information

		N
Cases	Valid	591
	Missing	19
	Weighted Valid	591
Cells	Defined Cells	128
	Structural Zeros	0
	Sampling Zeros	20
Categories	A	2
	FQuartile	4
	GenderX	2
	AgeBinary	2
	RevoveryD	2
	RecoveryA	2

Convergence Informationa,b

Maximum Number of Iterations	20
Converge Tolerance	.00100
Final Maximum Absolute Difference	6.42417E-7c
Final Maximum Relative Difference	6.15944E-7
Number of Iterations	4

a Model: Multinomial

b Design: Constant + A + FQuartile + GenderX + AgeBinary + RevoveryD + RecoveryA

c The iteration converged because the maximum absolute changes of parameter estimates is less than the specified convergence criterion.

Iteration Historyb,c

Iteration	Log Likelihood	Parameter									
		Constant	[A = 0]	[FQuartile = 1]	[FQuartile = 2]	[FQuartile = 3]	[GenderX = 0]	[AgeBinary = 0]	[RevoveryD = 0]	[RecoveryA = 0]	
0	904.103	1.5298	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	
1	983.637	2.1547	-.0372	-.1692	-.1286	-.1692	-.9577	.1387	-.1794	-.2335	
2	984.058	2.1739	-.0372	-.1643	-.1223	-.1643	-1.0413	.1390	-.1798	-.2346	
3	984.058	2.1743	-.0372	-.1643	-.1223	-.1643	-1.0430	.1390	-.1798	-.2346	
4	984.058a	2.1743	-.0372	-.1643	-.1223	-.1643	-1.0430	.1390	-.1798	-.2346	

Redundant parameters are not displayed. Their values are always zero in all iterations.

a The iteration converged because the maximum absolute changes of parameter estimates is less than the specified convergence criterion.

b Model: Multinomial
c Design: Constant + A + FQuartile + GenderX + AgeBinary + RevoveryD + RecoveryA

Goodness-of-Fit Testsa,b

	Value	df	Sig.
Likelihood Ratio	465.134	119	.000
Pearson Chi-Square	485.295	119	.000

a Model: Multinomial
b Design: Constant + A + FQuartile + GenderX + AgeBinary + RevoveryD + RecoveryA

Design Matrixa,b

Parameter A

0

1

FQuartile

FQuartile

1

2

3

4

1

2

3

4

GenderX

GenderX

GenderX

GenderX

GenderX

GenderX

GenderX

GenderX

GenderX

0 1

0

1 0

1

0 1

0

1 0

1

0 1
0
1
AgeBinary AgeBinary
AgeBinary
AgeBinary
AgeBinary AgeBinary
AgeBinary
AgeBinary
AgeBinary AgeBinary
AgeBinary
AgeBinary
AgeBinary AgeBinary
AgeBinary
AgeBinary
0 1 0
1 0 1
0 1 0
1 0 1
0 1 0
1 0 1
0 1 0
1 0 1
0 1 0
1 0 1
0 1
RevoveryD RevoveryD RevoveryD
RevoveryD RevoveryD
RevoveryD RevoveryD
RevoveryD RevoveryD RevoveryD
RevoveryD RevoveryD
RevoveryD RevoveryD
RevoveryD RevoveryD RevoveryD
RevoveryD RevoveryD
RevoveryD RevoveryD
RevoveryD RevoveryD RevoveryD
RevoveryD RevoveryD
RevoveryD RevoveryD
RevoveryD RevoveryD RevoveryD
RevoveryD
0 1 0 1 0 1
0 1 0 1 0 1
0 1 0 1 0 1
0 1 0 1 0 1
0 1 0 1 0 1
0 1 0 1 0 1
0 1 0 1 0 1

```
            0       1       0       1       0       1
            0       1       0       1       0       1
            0       1       0       1       0       1
            0       1       0       1
            RecoveryA   RecoveryA   RecoveryA   RecoveryA
            RecoveryA   RecoveryA   RecoveryA   RecoveryA
            RecoveryA   RecoveryA   RecoveryA   RecoveryA
            RecoveryA   RecoveryA   RecoveryA   RecoveryA
            RecoveryA   RecoveryA   RecoveryA   RecoveryA
            RecoveryA   RecoveryA   RecoveryA   RecoveryA
            RecoveryA   RecoveryA   RecoveryA   RecoveryA
            RecoveryA   RecoveryA   RecoveryA   RecoveryA
            RecoveryA   RecoveryA   RecoveryA   RecoveryA
            RecoveryA   RecoveryA   RecoveryA   RecoveryA
            RecoveryA   RecoveryA   RecoveryA   RecoveryA
            RecoveryA   RecoveryA   RecoveryA   RecoveryA
            RecoveryA   RecoveryA   RecoveryA   RecoveryA
            RecoveryA   RecoveryA   RecoveryA   RecoveryA
            RecoveryA   RecoveryA   RecoveryA   RecoveryA
            RecoveryA   RecoveryA   RecoveryA   RecoveryA
            0   1   0   1   0   1   0   1   0   1   0   1
            0   1   0   1   0   1   0   1   0   1   0   1
            0   1   0   1   0   1   0   1   0   1   0   1
            0   1   0   1   0   1   0   1   0   1   0   1
            0   1   0   1   0   1   0   1   0   1   0   1
            0   1   0   1   0   1   0   1   0   1   0   1
            0   1   0   1   0   1   0   1   0   1   0   1
            0   1   0   1   0   1   0   1   0   1   0   1
            0   1   0   1   0   1   0   1   0   1   0   1
            0   1   0   1   0   1   0   1   0   1   0   1
            0   1   0   1   0   1   0   1
Cell Structure  1   1   1   1   1   1   1   1   1   1   1
            1   1   1   1   1   1   1   1   1   1   1   1
            1   1   1   1   1   1   1   1   1   1   1   1
            1   1   1   1   1   1   1   1   1   1   1   1
            1   1   1   1   1   1   1   1   1   1   1   1
            1   1   1   1   1   1   1   1   1   1   1   1
            1   1   1   1   1   1   1   1   1   1   1   1
            1   1   1   1   1   1   1   1   1   1   1   1
            1   1   1   1   1   1   1   1   1   1   1   1
            1   1   1   1   1   1   1   1   1   1   1   1
            1   1   1   1   1   1   1   1   1
Constant        1   1   1   1   1   1   1   1   1   1   1
            1   1   1   1   1   1   1   1   1   1   1   1
            1   1   1   1   1   1   1   1   1   1   1   1
            1   1   1   1   1   1   1   1   1   1   1   1
            1   1   1   1   1   1   1   1   1   1   1   1
```

	1	1	1	1	1	1	1	1	1	1	1	1
	1	1	1	1	1	1	1	1	1	1	1	1
	1	1	1	1	1	1	1	1	1	1	1	1
	1	1	1	1	1	1	1	1	1	1	1	1
	1	1	1	1	1	1	1	1	1	1	1	1
	1	1	1	1	1	1	1	1	1			
[A = 0]	1	1	1	1	1	1	1	1	1	1	1	1
	1	1	1	1	1	1	1	1	1	1	1	1
	1	1	1	1	1	1	1	1	1	1	1	1
	1	1	1	1	1	1	1	1	1	1	1	1
	1	1	1	1	1	1	1	1	1	1	1	1
	1	1	1	1	0	0	0	0	0	0	0	0
	0	0	0	0	0	0	0	0	0	0	0	0
	0	0	0	0	0	0	0	0	0	0	0	0
	0	0	0	0	0	0	0	0	0	0	0	0
	0	0	0	0	0	0	0	0	0	0	0	0
	0	0	0	0	0	0	0	0				
[FQuartile = 1]	1	1	1	1	1	1	1	1	1	1	1	1
	1	1	1	1	1	0	0	0	0	0	0	0
	0	0	0	0	0	0	0	0	0	0	0	0
	0	0	0	0	0	0	0	0	0	0	0	0
	0	0	0	0	0	0	0	0	0	0	0	0
	0	0	0	0	0	1	1	1	1	1	1	1
	1	1	1	1	1	1	1	1	1	0	0	0
	0	0	0	0	0	0	0	0	0	0	0	0
	0	0	0	0	0	0	0	0	0	0	0	0
	0	0	0	0	0	0	0	0	0	0	0	0
	0	0	0	0	0	0	0	0	0			
[FQuartile = 2]	0	0	0	0	0	0	0	0	0	0	0	0
	0	0	0	0	0	1	1	1	1	1	1	1
	1	1	1	1	1	1	1	1	1	0	0	0
	0	0	0	0	0	0	0	0	0	0	0	0
	0	0	0	0	0	0	0	0	0	0	0	0
	0	0	0	0	0	0	0	0	0	0	0	0
	0	0	0	0	0	0	0	0	0	1	1	1
	1	1	1	1	1	1	1	1	1	1	1	1
	1	0	0	0	0	0	0	0	0	0	0	0
	0	0	0	0	0	0	0	0	0	0	0	0
	0	0	0	0	0	0	0	0	0			
[FQuartile = 3]	0	0	0	0	0	0	0	0	0	0	0	0
	0	0	0	0	0	0	0	0	0	0	0	0
	0	0	0	0	0	0	0	0	0	1	1	1
	1	1	1	1	1	1	1	1	1	1	1	1
	1	0	0	0	0	0	0	0	0	0	0	0
	0	0	0	0	0	0	0	0	0	0	0	0
	0	0	0	0	0	0	0	0	0	0	0	0
	0	0	0	0	0	0	0	0	0	0	0	0

```
        0       1       1       1       1       1       1       1       1       1       1       1
        1       1       1       1       1       0       0       0       0       0       0       0
        0       0       0       0       0       0       0       0       0
[GenderX = 0]   1       1       1       1       1       1       1       1       0       0       0
        0       0       0       0       0       1       1       1       1       1       1       1
        1       0       0       0       0       0       0       0       0       1       1       1
        1       1       1       1       1       0       0       0       0       0       0       0
        0       1       1       1       1       1       1       1       1       0       0       0
        0       0       0       0       0       1       1       1       1       1       1       1
        1       0       0       0       0       0       0       0       0       1       1       1
        1       1       1       1       1       0       0       0       0       0       0       0
        0       1       1       1       1       1       1       1       1       0       0       0
        0       0       0       0       0       1       1       1       1       1       1       1
        1       0       0       0       0       0       0       0       0
[AgeBinary = 0]         1       1       1       1       0       0       0       0       1       1
        1       1       0       0       0       0       1       1       1       1       0       0
        0       0       1       1       1       1       0       0       0       0       1       1
        1       1       0       0       0       0       1       1       1       1       0       0
        0       0       1       1       1       1       0       0       0       0       1       1
        1       1       0       0       0       0       1       1       1       1       0       0
        0       0       1       1       1       1       0       0       0       0       1       1
        1       1       0       0       0       0       1       1       1       1       0       0
        0       0       1       1       1       1       0       0       0       0       1       1
        1       1       0       0       0       0       1       1       1       1       0       0
        0       0       1       1       1       1       0       0       0       0
[RevoveryD = 0]         1       1       0       0       1       1       0       0       1       1
        0       0       1       1       0       0       1       1       0       0       1       1
        0       0       1       1       0       0       1       1       0       0       1       1
        0       0       1       1       0       0       1       1       0       0       1       1
        0       0       1       1       0       0       1       1       0       0       1       1
        0       0       1       1       0       0       1       1       0       0       1       1
        0       0       1       1       0       0       1       1       0       0       1       1
        0       0       1       1       0       0       1       1       0       0       1       1
        0       0       1       1       0       0       1       1       0       0       1       1
        0       0       1       1       0       0       1       1       0       0       1       1
        0       0       1       1       0       0       1       1       0       0
[RecoveryA = 0]         1       0       1       0       1       0       1       0       1       0
        1       0       1       0       1       0       1       0       1       0       1       0
        1       0       1       0       1       0       1       0       1       0       1       0
        1       0       1       0       1       0       1       0       1       0       1       0
        1       0       1       0       1       0       1       0       1       0       1       0
        1       0       1       0       1       0       1       0       1       0       1       0
        1       0       1       0       1       0       1       0       1       0       1       0
        1       0       1       0       1       0       1       0       1       0       1       0
        1       0       1       0       1       0       1       0       1       0       1       0
        1       0       1       0       1       0       1       0       1       0       1       0
        1       0       1       0       1       0       1       0       1       0
```

The default display of the design matrix is transposed. Redundant parameters are not displayed

a Model: Multinomial

b Design: Constant + A + FQuartile + GenderX + AgeBinary + RevoveryD + RecoveryA

Cell Counts and Residualsa,b

A	FQuartile	GenderX	AgeBinary	RevoveryD	RecoveryA	Observed		Expected		Residual		Standardized Residual		Adjusted Residual		Deviance	
						Count	%	Count	%								
0	1	0	0	0	0	11	1.9%	1.924	0.3%	9.076	6.554	6.669	6.193				

					1	1	0.2%	2.433	0.4%	-1.433	-.920	-.940	-1.333		
			1	0	3	0.5%	2.303	0.4%	.697	.460	.470	1.260			
					1	10	1.7%	2.912	0.5%	7.088	4.164	4.267	4.968		
		1	0	0	10	1.7%	1.674	0.3%	8.326	6.443	6.544	5.979			
					1	4	0.7%	2.117	0.4%	1.883	1.297	1.321	2.256		
			1	0	3	0.5%	2.004	0.3%	.996	.705	.717	1.556			
					1	4	0.7%	2.534	0.4%	1.466	.923	.943	1.911		
		1	0	0	0	21	3.6%	5.459	0.9%	15.541	6.682	6.944	7.522		
					1	5	0.8%	6.903	1.2%	-1.903	-.728	-.763	-1.796		
			1	0	6	1.0%	6.535	1.1%	-.535	-.210	-.220	-1.013			
					1	17	2.9%	8.263	1.4%	8.737	3.061	3.228	4.953		
		1	0	0	5	0.8%	4.751	0.8%	.249	.115	.119	.715			
					1	12	2.0%	6.007	1.0%	5.993	2.458	2.562	4.075		
			1	0	2	0.3%	5.687	1.0%	-3.687	-1.554	-1.617	-2.045			
					1	17	2.9%	7.191	1.2%	9.809	3.681	3.860	5.409		
		2	0	0	0	2	0.3%	2.006	0.3%	-.006	-.005	-.005	-.113		
					1	2	0.3%	2.537	0.4%	-.537	-.338	-.345	-.975		
			1	0	1	0.2%	2.402	0.4%	-1.402	-.906	-.925	-1.324			
					1	6	1.0%	3.037	0.5%	2.963	1.705	1.748	2.859		
		1	0	0	0	0.0%	1.746	0.3%	-1.746	-1.323	-1.344	.000			
					1	0	0.0%	2.208	0.4%	-2.208	-1.489	-1.518	.000		
			1	0	0	0.0%	2.090	0.4%	-2.090	-1.448	-1.475	.000			
					1	4	0.7%	2.643	0.4%	1.357	.837	.856	1.821		

		1	0	0	0	11	1.9%	5.693	1.0%	5.307	2.235	2.324	3.806		
					1	4	0.7%	7.199	1.2%	-3.199	-1.199	-1.257	-2.168		
				1	0	3	0.5%	6.815	1.2%	-3.815	-1.470	-1.538	-2.219		
					1	16	2.7%	8.617	1.5%	7.383	2.534	2.675	4.450		
			1	0	0	11	1.9%	4.955	0.8%	6.045	2.727	2.825	4.189		
					1	3	0.5%	6.265	1.1%	-3.265	-1.311	-1.368	-2.102		
				1	0	4	0.7%	5.931	1.0%	-1.931	-.797	-.830	-1.775		
					1	18	3.0%	7.499	1.3%	10.501	3.859	4.052	5.614		
		3	0	0	0	0	0	0.0%	1.924	0.3%	-1.924	-1.389	-1.414	.000	
					1	1	0.2%	2.433	0.4%	-1.433	-.920	-.940	-1.333		
				1	0	0	0.0%	2.303	0.4%	-2.303	-1.521	-1.551	.000		
					1	3	0.5%	2.912	0.5%	.088	.052	.053	.423		
			1	0	0	1	0.2%	1.674	0.3%	-.674	-.522	-.530	-1.015		
					1	1	0.2%	2.117	0.4%	-1.117	-.769	-.784	-1.225		
				1	0	1	0.2%	2.004	0.3%	-1.004	-.711	-.723	-1.179		
					1	3	0.5%	2.534	0.4%	.466	.293	.300	1.006		
		1	0	0	0	7	1.2%	5.459	0.9%	1.541	.662	.688	1.865		
					1	1	0.2%	6.903	1.2%	-5.903	-2.260	-2.367	-1.966		
				1	0	2	0.3%	6.535	1.1%	-4.535	-1.784	-1.865	-2.176		
					1	5	0.8%	8.263	1.4%	-3.263	-1.143	-1.205	-2.241		
			1	0	0	1	0.2%	4.751	0.8%	-3.751	-1.728	-1.789	-1.765		
					1	2	0.3%	6.007	1.0%	-4.007	-1.643	-1.713	-2.097		
				1	0	1	0.2%	5.687	1.0%	-4.687	-1.975	-2.055	-1.865		

1 9 1.5% 7.191 1.2% 1.809 .679
.712 2.010
4 0 0 0 0 0 0 0.0% 2.267 0.4% -2.267 -1.509 -
1.538 .000
1 0 0.0% 2.867 0.5% -2.867 -1.697 -
1.737 .000
1 0 0 0.0% 2.714 0.5% -2.714 -1.651 -
1.688 .000
1 1 0.2% 3.432 0.6% -2.432 -1.317 -
1.352 -1.570
1 0 0 1 0.2% 1.973 0.3% -.973 -.694 -
.706 -1.166
1 0 0.0% 2.495 0.4% -2.495 -1.583 -
1.616 .000
1 0 1 0.2% 2.362 0.4% -1.362 -.888 -
.906 -1.311
1 2 0.3% 2.987 0.5% -.987 -.572 -
.586 -1.266
1 0 0 0 7 1.2% 6.434 1.1% .566 .224
.234 1.086
1 3 0.5% 8.135 1.4% -5.135 -1.813 -
1.906 -2.447
1 0 3 0.5% 7.702 1.3% -4.702 -1.705 -
1.789 -2.379
1 6 1.0% 9.738 1.6% -3.738 -1.208 -
1.279 -2.411
1 0 0 4 0.7% 5.600 0.9% -1.600 -.679 -
.705 -1.640
1 2 0.3% 7.080 1.2% -5.080 -1.921 -
2.009 -2.249
1 0 1 0.2% 6.703 1.1% -5.703 -2.215 -
2.313 -1.951
1 5 0.8% 8.475 1.4% -3.475 -1.202 -
1.266 -2.297
1 1 0 0 0 0 0 0.0% 1.997 0.3% -1.997 -1.416 -
1.441 .000
1 0 0.0% 2.525 0.4% -2.525 -1.592 -
1.627 .000
1 0 1 0.2% 2.390 0.4% -1.390 -.901 -
.920 -1.320
1 0 0.0% 3.022 0.5% -3.022 -1.743 -
1.787 .000
1 0 0 0 0.0% 1.738 0.3% -1.738 -1.320 -
1.341 .000
1 1 0.2% 2.197 0.4% -1.197 -.809 -
.825 -1.255

				1	0	0	0.0%	2.080	0.4%	-2.080	-1.445	-1.472	.000		
					1	0	0.0%	2.630	0.4%	-2.630	-1.625	-1.662	.000		
		1	0	0	0	0	0.0%	5.667	1.0%	-5.667	-2.392	-2.489	.000		
					1	1	0.2%	7.165	1.2%	-6.165	-2.317	-2.430	-1.985		
				1	0	0	0.0%	6.783	1.1%	-6.783	-2.619	-2.742	.000		
					1	4	0.7%	8.576	1.5%	-4.576	-1.574	-1.663	-2.470		
			1	0	0	1	0.2%	4.931	0.8%	-3.931	-1.778	-1.842	-1.786		
					1	0	0.0%	6.235	1.1%	-6.235	-2.510	-2.620	.000		
				1	0	0	0.0%	5.903	1.0%	-5.903	-2.442	-2.544	.000		
					1	1	0.2%	7.463	1.3%	-6.463	-2.381	-2.501	-2.005		
	2	0	0	0	0	4	0.7%	2.082	0.4%	1.918	1.331	1.356	2.285		
					1	3	0.5%	2.633	0.4%	.367	.227	.232	.885		
				1	0	3	0.5%	2.493	0.4%	.507	.322	.329	1.054		
					1	3	0.5%	3.152	0.5%	-.152	-.086	-.088	-.544		
			1	0	0	1	0.2%	1.812	0.3%	-.812	-.604	-.614	-1.090		
					1	1	0.2%	2.291	0.4%	-1.291	-.855	-.872	-1.288		
				1	0	1	0.2%	2.169	0.4%	-1.169	-.795	-.811	-1.245		
					1	1	0.2%	2.743	0.5%	-1.743	-1.055	-1.079	-1.421		
		1	0	0	0	11	1.9%	5.909	1.0%	5.091	2.105	2.192	3.697		
					1	2	0.3%	7.472	1.3%	-5.472	-2.015	-2.115	-2.296		
				1	0	5	0.8%	7.074	1.2%	-2.074	-.784	-.822	-1.863		
					1	6	1.0%	8.944	1.5%	-2.944	-.992	-1.049	-2.189		
			1	0	0	8	1.4%	5.143	0.9%	2.857	1.266	1.312	2.659		

						1	3	0.5%	6.502	1.1%	-3.502	-1.381	-1.443	-2.154	
				1		0	0	0.0%	6.156	1.0%	-6.156	-2.494	-2.600	.000	
						1	9	1.5%	7.783	1.3%	1.217	.439	.462	1.617	
3	0	0	0	0		0	4	0.7%	1.997	0.3%	2.003	1.420	1.445	2.357	
						1	1	0.2%	2.525	0.4%	-1.525	-.962	-.983	-1.361	
				1		0	3	0.5%	2.390	0.4%	.610	.395	.403	1.167	
						1	4	0.7%	3.022	0.5%	.978	.564	.578	1.497	
			1	0		0	3	0.5%	1.738	0.3%	1.262	.959	.974	1.810	
						1	2	0.3%	2.197	0.4%	-.197	-.133	-.136	-.613	
				1		0	4	0.7%	2.080	0.4%	1.920	1.333	1.358	2.287	
						1	5	0.8%	2.630	0.4%	2.370	1.465	1.498	2.535	
1			0	0		0	12	2.0%	5.667	1.0%	6.333	2.673	2.782	4.244	
						1	5	0.8%	7.165	1.2%	-2.165	-.814	-.853	-1.897	
				1		0	2	0.3%	6.783	1.1%	-4.783	-1.847	-1.933	-2.210	
						1	19	3.2%	8.576	1.5%	10.424	3.586	3.787	5.498	
			1	0		0	8	1.4%	4.931	0.8%	3.069	1.388	1.438	2.782	
						1	8	1.4%	6.235	1.1%	1.765	.711	.742	1.997	
				1		0	5	0.8%	5.903	1.0%	-.903	-.374	-.389	-1.288	
						1	17	2.9%	7.463	1.3%	9.537	3.513	3.690	5.290	
4	0	0	0	0		0	5	0.8%	2.353	0.4%	2.647	1.729	1.763	2.745	
						1	3	0.5%	2.976	0.5%	.024	.014	.014	.221	
				1		0	5	0.8%	2.817	0.5%	2.183	1.304	1.334	2.395	
						1	7	1.2%	3.562	0.6%	3.438	1.827	1.878	3.075	

			1	0	0	7	1.2%	2.048	0.3%	4.952	3.466	3.528	4.148		
					1	1	0.2%	2.590	0.4%	-1.590	-.990	-1.011	-1.379		
				1	0	2	0.3%	2.452	0.4%	-.452	-.289	-.295	-.902		
					1	3	0.5%	3.100	0.5%	-.100	-.057	-.058	-.443		
		1	0	0	0	16	2.7%	6.678	1.1%	9.322	3.628	3.787	5.288		
					1	7	1.2%	8.444	1.4%	-1.444	-.501	-.527	-1.620		
				1	0	7	1.2%	7.994	1.4%	-.994	-.354	-.372	-1.364		
					1	15	2.5%	10.108	1.7%	4.892	1.552	1.646	3.441		
			1	0	0	11	1.9%	5.812	1.0%	5.188	2.163	2.248	3.746		
					1	7	1.2%	7.348	1.2%	-.348	-.129	-.135	-.825		
				1	0	9	1.5%	6.957	1.2%	2.043	.779	.815	2.153		
					1	24	4.1%	8.796	1.5%	15.204	5.165	5.447	6.941		

a Model: Multinomial

b Design: Constant + A + FQuartile + GenderX + AgeBinary + RevoveryD + RecoveryA