Evaluation of the Brief Addiction Therapist Scale (BATS) through the secondary analysis of data

Kate Thresher

Submitted in accordance with the requirements for the degree of

Doctor of Clinical Psychology (D. Clin. Psychol.)

The University of Leeds

School of Medicine

Division of Psychological and Social Medicine

June 2020

The candidate confirms that the work submitted is her own and that appropriate credit has been given where reference has been made to the work of others.

Acknowledgements

Firstly, thank you to my supervisors, Drs Gary Latchford, Bridgette Bewick and Helen Crosby for their unwavering enthusiasm, support and thought-provoking supervision. You have been fantastic supervisors and it has been a pleasure to work on this project with you. I would also like to extend my thanks to Dr Rebecca Walwyn for providing her statistical expertise on my data. Thank you all for sharing your research wisdom and for guiding me through this process.

Thank you also to the teams who provided me with access to the data sets, and to the participants in UKATT who consented to their data being used in secondary research, without which this research would not have been possible.

To the trainees in my cohort, thank you all for three wonderful years of some of the best fun, constant support and encouragement. I am so proud of us all for what we have achieved and feel extremely lucky to have shared my DClin journey with you.

Finally, to my family, friends outside of training and partner. Thank you all for your kindness, support and unconditional love. Thank you for listening when I told you that I couldn't do it and for continuing to believe in me anyway. Your ongoing encouragement has spurred me on throughout. Charlie – thank you for always being there, knowing what to say and how to make me laugh during the challenges of thesis life!

Abstract

The aim of this research study was to investigate whether the Brief Addiction Therapist Scale (BATS) (Crosby, 2018), a transtheoretical fidelity measure, is associated with therapeutic outcome. The BATS evaluates therapists' routine delivery of psychotherapies widely used in the treatment of alcohol and drug use problems. The current research conducted a secondary analysis of existing randomised controlled trial (RCT) data. Digital recordings of Motivational Enhancement Therapy (MET) sessions that were recorded as part of the original United Kingdom Alcohol Treatment Trial (UKATT Team, 2005) were observed and therapists were rated using the BATS. Multilevel statistical procedures in which the therapists were treated as a random factor were conducted in order to examine whether the BATS scores could predict therapy outcome. The results showed that the therapist BATS scores did not predict clients' outcome of therapy. However, the analysis suggested that the BATS scores explain some of the therapist variance in outcomes. The findings of this thesis provide further support for the real-world application of the BATS.

Contents

List of Tables

List of Figures

List of Abbreviations

A-CRA	Adolescent Community Reinforcement Approach
ADAPTA	Addressing Drinking Among Patients: comparing Two Approaches
AESOPS PRS	Alcohol: Evaluating Stepped care in Older Populations Study Process Rating Scale
ASI	Addiction Severity Index
BATS	Brief Addiction Therapist Scale
BDI	Beck Depression Inventory
CBT	Cognitive Behavioural Therapy
CI	Confidence Interval
CM	Contingency Management
CPT	Cognitive Processing Therapy
DDD	Drinks per Drinking Day
ICC	Interclass Correlation Coefficient
IDC	Individual Drug Counselling
IRR	Inter-Rater Reliability
ITRS	Independent Tape Rater Scale
MATCH	Matching Alcohol Treatments to Client Homogeneity
MET	Motivational Enhancement Therapy
MI	Motivational Interviewing
MITI	Motivational Interviewing Treatment Integrity Scale
NHS	National Health Service
NIHM TDCRP	National Institute of Mental Health Treatment of Depression Collaborative Research Program
PI	Principal Investigator

PRS	Process Rating Scale
PTSD	Post-Traumatic Stress Disorder
RCT	Randomised Controlled Trial
SBNT	Social Behaviour and Network Therapy
UKATT	United Kingdom Alcohol Treatment Trial

Chapter One: Introduction

1.1 Introduction

The main aim of my thesis was to investigate whether the Brief Addiction Therapist Scale (BATS) (Appendix A), a transtheoretical fidelity measure for evaluating therapists' delivery of psychological therapies for alcohol and drug use problems, is associated with therapeutic outcome. This chapter introduces the background literature to the field of addictions and its treatment, the research on what potentially makes psychotherapy effective, and the role of fidelity measurement. A summary of a systematic literature search exploring studies that have examined the relationship between fidelity measurement and therapy outcome is discussed. The chapter concludes with an introduction to the BATS measure and the research question for my thesis project.

1.2 Background

The World Health Organisation (2018) report that harmful use of alcohol accounts for around 5.3% of deaths worldwide. Overall, 5.1% of the global burden of disease and injury is attributed to alcohol. In 2017/2018, there were around 338,000 estimated hospital admissions in England whereby the main reason for admission was related to alcohol. This figure represented a 15% increase compared to the number of admissions in 2007/2008 (NHS Digital, 2019). There is also a large amount of evidence regarding the adverse health effects of drug use. For instance, opioid dependence is considered to be a substantial contributor to the global burden of disease (Degenhardt et al., 2014) and cannabis use has been shown to

have a positive association with anxiety disorders (Kedzior & Tabata Laeber, 2014). The widespread problems with alcohol and drug dependence led researchers and clinicians to attempt to identify effective therapeutic treatments. Psychological therapies found to be effective and thus widely used for the treatment of substance use and dependence include: Motivational Enhancement Therapy (MET), Cognitive Behavioural Therapy (CBT), and Social Behaviour and Network Therapy (SBNT) (Miller et al., 1992; Carroll, 1998; Copello et al., 2002).

1.3 What makes therapy effective?

The effectiveness and efficacy of psychotherapy treatments is well established across a range of treatment models and presenting conditions. For instance, CBT, the most researched form of psychotherapy, is frequently referred to as the "gold-standard" psychological treatment. To date, no other form of psychotherapy has demonstrated evidence of being systematically superior to CBT. Consequently, CBT dominates the guidelines for psychosocial interventions, making this a first-line treatment for many mental health difficulties (David et al., 2018). Whilst research into the efficacy of psychological therapies has been important, the process of ascertaining how psychotherapy works and what makes it effective has been more problematic. Consequently, researchers have proceeded to attempt to identify and explore how therapy models generate change in clients.

There has been some exploration as to whether the outcomes from differing psychotherapies are equivalent (Smith & Glass, 1977; Shapiro & Shapiro, 1982). Researchers have confirmed the hypothesis that if there are differences between the outcomes of psychotherapies then these are relatively small (Wampold & Imel, 2015). If these differences are in fact small and

therapy models have comparable outcomes, it is plausible that there are other factors besides those that are therapy specific, that influence the outcomes seen in day-to-day psychotherapy practice. Whilst acknowledging that there is also the potential that therapy models have distinct mechanisms that are equally effective, the above hypothesis highlights the possibility that the effectiveness of therapy may be due to factors which are common across therapeutic models. Researchers have therefore suggested that specific theory-based interventions used in therapy models are likely to be less important than the "common factors" shared across models. It is these "common factors" that are hypothesised to be responsible for a large portion of the change attributable to the intervention in therapy (Sprenkle & Blow, 2004; Wampold & Imel, 2015).

The term "common factors" was coined by Rosenzweig (1936) who was the first to propose the idea that the mechanisms of change in therapy may not be due to specific therapy models but instead may reflect factors shared across them. Rosenzweig (1936) made some suggestions regarding "common factors" that may better explain therapy success. These included: a) therapist qualities such as inspiring aspects of the therapist's personality, b) the therapeutic relationship facilitating psychological processes such as social reconditioning or catharsis, c) the reformulation of psychological events, and d) the reintegration of personality through the systematic application of a therapeutic ideology. However, Rosenzweig's paper is theoretical and contains no data to support his arguments.

Since then, "common factors" in psychotherapy have been increasingly examined with the majority of research investigating those relating to: (a) therapist qualities, such as empathy, (b) client qualities e.g. expectations, and (c) the therapeutic relationship. One particular "common factor" model is the Contextual Model (Wampold & Imel, 2015) which suggests

that there are three ways in which psychotherapy results in change. This suggests that psychotherapy does not have a single influence on a client but instead works through various mechanisms. These processes include the therapeutic alliance, the client's expectations, and specific elements of the particular therapy (e.g. therapeutic techniques and actions). The Contextual Model provides an explanation for the effects of psychotherapy that is an alternative to those which suggest specific ingredients of a therapy model are required and are beneficial for specific disorders (Wampold & Imel, 2015).

Research findings indicate that most therapeutic models share some common aspects (Sprenkle et al., 2013) and that these are related to outcome (Wampold & Imel, 2015). Lambert & Barley (2002) concluded in their summary of the psychotherapy literature, that common factors accounted for 30% of the variance in treatment outcome, extra-therapeutic factors 40%, specific techniques 15% and placebo or expectation effects for 15%. Research by Wampold (2001) suggested that client factors (unique aspects of each client and his or her environment) and extra therapy events (events occurring outside the treatment) accounted for the vast majority of variance in psychotherapy outcome (87%). In addition to this, of the proportion accounted for by the specific therapy (13%), the therapeutic relationship was found to be the most important factor, further emphasising the importance of the "common factors". The research clearly suggests that common factors account for a large portion of variance in therapeutic outcomes (between 30% and 87%) and are therefore likely to be important variables in the psychotherapy process.

What has been clearly highlighted within the literature is that the therapist is an important factor relating to psychotherapy outcomes. The therapists who generally achieve more positive outcomes are those that: possess the ability to form a good therapeutic alliance with a

range of different clients, that show interest in their own therapeutic effectiveness and make concerted efforts to improve their practice (Wampold & Imel, 2015). Additionally, a purposeful collaborative therapeutic relationship between therapist and client has also been identified as being related to positive treatment outcome (e.g. Meier et al., 2005; Black et al., 2005), emphasising the importance of joint working in therapy towards achieving the client's goal.

It is recognised that in order to understand what makes therapy effective, we require ways of identifying 'active ingredients', and this relies on having accurate ways of measuring the different components of therapy. If we can measure these factors then this has the potential to inform and improve clinical practice, and potentially therapy outcomes. Emphasis is often placed on generating evidence for the efficacy of psychotherapy treatments as this is vital in developing the evidence base and informing routine clinical practice. Consequently, adequate fidelity measurement is required to enable us to understand the processes of therapy, as delivered by a therapist, and the effect that these can have for clients (Schoenwald et al., 2011). Having adequate fidelity measurement can enable a particular therapy under study to be specified and its outcomes demonstrated. This addresses the need to make accurate interpretations of therapy effects so that we can externalise the findings from a research study.

1.4 Fidelity measurement

There has been an increasing emphasis on the importance of using fidelity measurement in both research (Gearing et al., 2011), and in clinical practice settings. Fidelity measurement refers to the extent to which interventions are implemented as intended (Yeaton & Sechrest,

1981) and is a strategy to monitor and improve the reliability and validity of an intervention (Borrelli et al., 2005). Specifically, fidelity measurement is "the extent to which the user's current practice matches the … "ideal" (Loucks, 1983, p. 4). Using measures of fidelity allows us to examine the magnitude to which the core components of treatment are delivered in line with treatment protocols (Moncher & Prinz, 1991).

Fidelity measurement emerged in the research literature in the 1980s. Yeaton and Sechrest (1981) argued that strength, integrity and effectiveness were three important dimensions of treatment fidelity which should be considered when deciding upon treatments. It was later suggested that the use of treatment manuals was vital in ensuring that the delivery of therapies and their associated techniques was consistent (Waltz et al., 1993). In addition to this, it was proposed that successful measures of treatment fidelity not only measured adherence to treatment protocols, but also determined whether interventions were being delivered competently by therapists (Waltz, et al., 1993).

There is an increasing awareness that the term "fidelity" is complex and multi-faceted (Prowse & Nagel, 2015) and there appears to be a lack of clarity and consistency in the definition and construct of fidelity (Gearing et al., 2011). In research literature, there are a variety of terms for fidelity which are used interchangeably, these include, "treatment integrity", and "treatment fidelity". However, within psychotherapy research most definitions refer to three core components of treatment fidelity: treatment adherence, therapist competence, and treatment differentiation (Schoenwald & Garland, 2013), with the majority emphasising adherence and competence. In their meta-evaluation, Prowse and Nagel (2015) concluded that there were inconsistencies in the importance that researchers assigned to defining treatment fidelity in their research. They argued that a lack of a consistent

framework or treatment fidelity model perpetuates differences in the fidelity definition. This also raises questions around how fidelity is operationalised.

Therapist adherence is considered a critical component of fidelity and refers to the extent to which the evidence-based treatment is delivered as described in the treatment manual (Yeaton & Sechrest, 1981) i.e. "did the therapy occur has intended?" (Hogue et al., 1996, p. 335). It is the degree to which specific interventions and approaches of a therapy model are present, and the specific ingredients of other treatment models are absent. Adherence is considered a prerequisite for competence, as it is judged impossible to competently deliver a specific intervention without demonstrating adherence to it (Perepletchikova & Kazdin, 2005). This suggests that competence therefore requires some form of conforming to specific treatment guidelines. Adherence has been conceptualised as consisting of treatment behaviours that are i) unique and essential, ii) essential but not unique, iii) neither unique or essential, and iv) proscribed (Waltz et al., 1993). Having high levels of the first two behaviours would be described as good adherence (Barber et al., 2007). Measures of adherence have been considered essential in establishing experimental validity (Rapley, & Loades, 2019). However, whilst the theoretical importance of therapist adherence in psychotherapy has been discussed within the literature, its impact on treatment outcome remains ambiguous (Emmelkamp et al., 2014) as will be discussed further.

There is a lack of agreement about what is meant by therapist competence and how best to measure this (Elkin, 1999). Barber et al. (2007) suggest that there is a crucial distinction between adherence and competence. Adherence shows that one knows how to intervene and has the ability to do so, demonstrating context-independent knowledge. In contrast, competence is referred to as context-dependent knowledge, which requires the ability to

know when and when not to intervene. Commonly, therapist competence is described as the skill demonstrated by the clinician when implementing the intervention in question (Waltz et al., 1993). This can also include non-specific treatment effects such as, the ability for the therapist to engage with the client (Santacroce et al., 2004), therapist warmth (Borelli et al., 2005), and the sensitivity in which the interventions are delivered (Perepletchikova et al., 2007). Many researchers assess therapist competence by examining whether specific intervention techniques are delivered competently. There appears to be less focus on general therapist skills and "global competency" (i.e. the possession of clinical acumen). This is potentially due to the need to assess competency when delivering specific therapeutic interventions particularly in research trials, in addition to the challenges in attempting to define and measure global competency skills. There are differing levels of competence which adds to complexity in measuring this concept. Competence within a Randomised Controlled Trial (RCT) may have a different meaning compared to the concept of competence within general practice (Barber et al., 2007). Therefore, assessing competence for different purposes and within different contexts makes competence a difficult concept to measure.

Treatment differentiation refers to whether the treatments that are under investigation differ from each other along critical dimensions (Perepletchikova & Kazdin, 2005). It involves comparing how and the extent to which key components were delivered (e.g. competence, and adherence) within the interventions being compared. Treatment differentiation is unique in that it integrates adherence and competence, and incorporates a comparison between treatments (Hagermoser Sanetti et al., 2016). For instance, within research, treatment differentiation data is frequently used to ensure that a treatment being investigated is distinct from the control condition (Dane & Schneider, 1998). Whilst adherence and treatment differentiation are closely related in that adherence can enable us to determine whether

treatments are different (Waltz et al., 1993), there has been conflicting views around the relationship between treatment adherence and therapist competence. In many instances adherence and competence may exist or occur at the same time, making it difficult to distinguish between them. Indeed, some studies have found moderate to high correlations (ranging from r =0.5 to r = 0.9) between adherence and competence (e.g. Barber et al., 2003; Carroll et al., 2000; McDonnell et al., 2007) highlighting a difficulty in disentangling the two constructs. This has led to some researchers arguing that in order to be competent in therapy there needs to be adherence to the therapy modality protocol. They suggest therapist adherence is a prerequisite for competent delivery of therapy, but adherence provides no guarantee of competence (Waltz et al., 1993). This means that therapist adherence may be necessary but not sufficient for reaching competence (Rapley., & Loades, 2019). In contrast, other studies have concluded that adherence and competence are not highly related (e.g. Miller & Binder, 2002). Each component of treatment fidelity captures a unique aspect of fidelity that collaboratively, and/or in isolation may be responsible for therapeutic outcome (Perepletchikova et al., 2007).

1.5 The importance of using fidelity measures

Fidelity measures are considered vital in enabling researchers to develop accurate and meaningful interpretations of intervention effects, particularly within research studies such as RCTs, which examine the efficacy of treatment models. A failure in either treatment adherence or competence of the therapist may compromise treatment fidelity and threaten the internal validity of a study (Perepletchikova & Kazdin, 2005). For example, if therapists deliver treatment interventions not prescribed in the treatment manual then conclusions about the results for the target therapy would be ambiguous. Additionally, if therapists adhere to the

treatment specific techniques but do so with limited skill, then this may contribute to poor therapy outcomes which are ascribed to the therapy rather than its delivery. Treatment fidelity is therefore important in enabling us to evaluate the efficacy of treatments, as it allows for treatment effects to be accurately attributed (Tober et al., 2008). The enhancement of treatment fidelity can serve to not only increase internal validity but can also increase external validity, as fidelity is required for both replication and generalisation of interventions to applied settings (Borrelli et al., 2005). However, particularly within RCTs, there remains a lack of evidence for effective measurement of treatment adherence, competence and differentiation. Perepletchikova et al. (2007) identified in their review of RCTs of youth and adult psychotherapies that only 3.5% of those reviewed adequately conducted and reported treatment fidelity procedures.

In order to adequately assess fidelity, we need reliable and valid measures to monitor therapist delivery of treatments. In the addiction field, one of the earliest measures of fidelity to be developed for research purposes was the MATCH (Matching Alcohol Treatments to Client Homogeneity) Tape Rating Scale (Carroll et al., 1998). This was developed for use within Project MATCH, the largest study of psychological therapies for alcohol use problems (Project MATCH Research Group, 1997), to assess adherence and differentiate between, Cognitive Behavioural Coping Skills Therapy, Motivational Enhancement Therapy (MET) and Twelve-Step Facilitation Therapy. The MATCH Tape Rating Scale, enabled researchers to demonstrate that treatments were distinguishable from each other when interventions were delivered by therapists as prescribed in the treatment manuals (Carroll et al., 1998). This measure only evaluated two aspects of fidelity, adherence and differentiation, as therapist competence was not measured.

Measures of fidelity have traditionally been developed within the context of research trials but it has been suggested that there is a role for these measures within clinical practice where the use of evidence-based treatments is considered crucial (Manuel, Hagedorn, & Finney, 2011), and fidelity measurement can potentially ensure that these evidence-based treatments are delivered appropriately (Miller et al., 2005). There is a growing need for improvement in the use of fidelity measurement methods in routine care (Schoenwald & Garland, 2013). Measurement that monitors the delivery of therapies in routine practice is necessary for guiding successful implementation (Schoenwald & Garland, 2013). Furthermore, indicators of fidelity are essential for stakeholders in mental health settings (e.g. commissioners, clients, practitioners) to determine whether clients' psychotherapy outcomes are attributable to a particular treatment or, of its application (Schoenwald, et al., 2011). For example, a therapy may be discontinued because it is thought to be ineffective when in fact it was the way that the therapy was delivered that resulted in it being ineffective. Different stakeholders may monitor adherence for a number of reasons. For example, service providers that are contracted and accountable for delivering a specific treatment, may use adherence data to provide evidence that this treatment was delivered.

Through the assessment of therapists' adherence to treatment models and their competence to deliver such interventions, information regarding individual therapist training needs may be identified. Fidelity measurement therefore has the potential to inform clinical supervision and clinician development needs. Currently the assessment and monitoring of a therapist's fidelity to a treatment model in clinical practice mostly relies on therapist self-report. Therapists tend to be overly positive in their evaluations of their own adherence and competence to therapy models (Breitenstein et al., 2010; Martino et al., 2008). In support of this observation, Carroll et al. (2010) argue that self-report should not be solely relied on to

monitor treatment fidelity in routine clinical practice. While video or audio recordings of therapy sessions are occasionally used within clinical supervision in order to gather feedback on therapist skill, these are rarely used alongside a fidelity measure, which would enable the formal evaluation of therapy delivery by the observer. Treatment fidelity measures are therefore required to support implementation and delivery of evidence-based practice. In addition to the utility of measures that assess fidelity to specific therapy models, there is the potential for measures that assess fidelity to good therapeutic practice to be valuable in clinical practice, particularly considering the common factors research. For instance, the importance of a therapist displaying empathy to their clients. Indeed, some of the earliest measurement scales to assess therapist empathy include the Truaz-Carhuff Accurate Empathy Scale (Truax & Carkhuff, 1967) which has demonstrated evidence as a reliable predictor of positive outcomes in therapy (Truax & Carkhuff, 1967).

Through using fidelity measures in a clinical context, supervisors can gather important information on the strengths and areas of difficulty for each therapist (e.g. by reviewing a therapist's session with a client and providing this feedback) (Bassett, Stein, Rossi & Martin, 2016). Rating therapists' sessions using a fidelity measure can also provide information to improve the quality of therapist training and ultimately client treatment (Waltz et al., 1993). It is possible that through the measurement of fidelity, there is the potential to develop therapist skill, further increasing the possibility of a positive client response to treatment. However, to effectively monitor fidelity of an intervention in routine care, measurement methods need to be ecologically valid, and feasible (Manderscheid, 1998; Hayes, 1998). This is a particular challenge in itself as such measures may encompass two separate constructs, which are therapist adherence and therapist competence (Perepletchikova et al., 2007).

1.6 Fidelity and Treatment Outcome

An underlying assumption of treatment fidelity measurement is that if therapist fidelity is maintained then it should relate to therapeutic change (Perepletchikova & Kazdin, 2005). However, researchers thus far have been unable to identify strong links between therapists' fidelity to treatment protocols and client outcomes (Miller & Binder, 2002), and there are conflicting findings within the literature.

I completed a systematic literature search using databases PsycINFO and Medline from the earliest available date to November 2019 (PsychINFO 1806 to 2019 and Medline 1946 to 2019), to identify studies which had examined the relationship between treatment fidelity (adherence and competence) and therapy outcome. The search strategy utilised can be viewed in Appendix B. Other potentially relevant articles were identified by hand searching reference lists of articles retrieved from the database searches. Together these search strategies yielded a total of 404 papers.

From this initial search, I identified two reviews that had examined studies investigating the relationship between therapist fidelity and therapy outcome (Perepletchikova & Kazdin, 2005; Webb et al., 2010), both of which had revealed mixed findings. Perepletchikova and Kazdin (2005) found that a number of studies reported a positive correlation between adherence and therapy outcome. However, a similar number of studies, including a large study into the treatment of depression (Elkin, 1999), identified no relationship between fidelity and treatment outcome. In the most recent meta-analysis involving thirty-six studies investigating the role of therapist adherence and competence in relation to treatment outcome, Webb et al. (2010) found that there were inconsistencies amongst the studies that they

reviewed. When combined across different treatment modalities for a range of presenting conditions, the meta-analysis indicated that effect sizes were small and nonsignificant for the mean weighted therapist adherence and intervention outcome association (r = 0.02 based on thirty-two studies) and the therapist competence-outcome relationship (r = 0.07 based on seventeen studies).

Based on the findings of the reviews described above, I adjusted my search strategy to focus on capturing research conducted after these reviews. I searched the databases from January 2008 to November 2019 which yielded a total of 257 papers. Papers were then selected for my literature review based on the following inclusion and exclusion criteria:

Inclusion criteria

- The paper had been published between January 2008 and November 2019
- Therapists in the study delivered an individual psychotherapy intervention
- Therapist adherence and/or competence had been measured
- The researchers examined the relationship between therapist adherence and/or competence and treatment outcome

Exclusion criteria

- Papers that had been published prior to January 2008
- The treatment intervention in the study was not delivered on an individual basis e.g. group psychotherapy
- Studies that had not measured therapist adherence and/or competence
- Studies that had not examined the relationship between adherence and/or competence ratings and outcome of therapy

Based on the inclusion and exclusion criteria fifteen papers were considered relevant as they had directly investigated the relationship between therapist fidelity and therapy outcome in individual psychotherapies; these are summarised in Appendix C. I read each paper and identified what elements of therapist fidelity had been measured by the researchers, i.e. adherence and/or competence, and whether they had identified a relationship between fidelity and therapy outcome. I then extracted this information to ascertain whether there was a common direction.

Overall my search of the literature published since 2008 identified mixed findings regarding the relationship between fidelity and treatment outcome, which was consistent with the two previous reviews (Perepletchikova & Kazdin, 2005; Webb et al., 2010). There were also inconsistencies as to whether researchers defined fidelity in their papers, with some referring either to adherence, competence or both adherence and competence, making this a complex concept to search. This potentially reflects earlier discussions around the lack of consistency in the definition and construct of fidelity (e.g. Gearing et al., 2011). As a result of these inconsistencies, the findings of my literature search are discussed in more detail below with reference to what aspects of fidelity were being measured in the study.

Some of the literature has identified that decreased treatment fidelity demonstrated by therapists can lead to decreased therapeutic change (e.g. Frank et al., 1991; Strunk et al., 2010). For example, Holder et al. (2018) explored the role of treatment fidelity (adherence and competence) on the outcomes of Cognitive Processing Therapy (CPT) for Post-Traumatic Stress Disorder (PTSD). Findings indicated that the therapists who demonstrated 'good' fidelity to the treatment model achieved better therapy outcomes. Specifically, these therapists' clients showed greater reductions in PTSD symptoms, trauma-related cognitions

and depression symptomatology, when compared to the clients of therapists who displayed 'below average' treatment fidelity. However, the inclusion of only four therapists delivering the treatment limited the researchers' ability to investigate and control for a wide variety of potential therapist factors that can impact treatment outcomes. Furthermore, the fidelity ratings were made by one independent rater who assigned therapists with a dichotomous variable for 'good' and 'below average' treatment fidelity, which removed variance in each therapist's rated fidelity.

There has been some suggestion that the timing of therapist adherence in stages of therapy is important for client outcome (Folke et al., 2017). Folke et al. (2017) investigated change in therapist adherence to CBT for bulimia nervosa over time and the relationship between adherence and outcome in the early (session three), middle (session eleven) and late phases of treatment (session twenty). Higher levels of therapist adherence in the early and middle phases of therapy (as measured by the Cognitive Behavioural Therapy Treatment Protocol Adherence Scale; Loeb et al., 2005) were associated with reduced frequency of binge eating. Higher levels of adherence measured in the later phases of treatment were not related to outcome. Similarly, Haug et al. (2016) reported that higher therapist fidelity (both adherence and competence) early in therapy was associated with more positive outcomes amongst clients presenting with panic disorder. Whilst both of these studies highlighted the importance of fidelity in the earlier stages of therapy, in a review of research investigating the efficacy of Multidimensional Family Therapy, Liddle (2010) concluded that when therapists generally drift away from treatment protocols during the process of therapy, the outcomes for clients are likely to deteriorate.

A number of studies do not support the relationship between treatment adherence and outcome (e.g. Weisman et al., 2002; Farmer et al., 2017; Brown et al., 2013; Weck et al., 2012). For instance, in a multicentre CBT outcome trial for the treatment of panic disorders, Huppert et al. (2001) identified that although all of the therapists satisfied the study's criteria for adherence, there was considerable variability in the outcomes reported by clients. Researchers found that the effect sizes for individual therapists ranged from 0-18%. This suggests that whilst some therapists were highly effective, some made minimal difference to their clients, despite all of the therapists delivering the intervention according to the tightly controlled treatment protocol. However, the researchers did not examine the direct relationship between therapist adherence and outcome, and it was unclear as to whether there were patient characteristics such as motivation, that interacted with adherence and outcome, as these were not explored. Interestingly, therapists were reported to differ in terms of their competence (as measured by a single outcome). However, there was no relationship found between competence and outcome.

Since then, Huppert, et al. (2006) have examined the relationship between patient motivation, therapist adherence and treatment outcomes for clients that were treated with CBT for panic disorder. Greater therapist protocol adherence, as measured by a scale developed specifically for the study, was associated with poorer outcome for clients that were deemed to be less motivated. Amongst the clients that were rated as more motivated, therapist adherence was not significantly associated with therapy outcome. Similarly, Hauke et al. (2014) identified that whilst adherence did not influence outcome at a global level, the clients' symptom severity and client motivation interacted with therapist adherence to predict outcome, emphasising the potential impact of other factors on the adherence and outcome relationship.

Therapists that are classed as being competent in their practice, could be expected to have better treatment outcomes than therapists considered to not be competent. As previously discussed, the complexities around defining and measuring competence have been highlighted in the literature. However, if we assume that measures used in research studies have captured what it means to be a competent therapist, does competence lead to better therapy outcomes? Indeed, a number of studies have explored the relationship between therapist competence and therapy outcome. Some have identified a positive relationship (e.g. Brown et al., 2013, O'Malley et al., 1988; Trepka et al., 2004). For instance, Haug et al. (2016) examined the association between alliance, adherence, competence and therapeutic outcomes for CBT for social anxiety and panic disorder. It was identified that higher therapist competence (and adherence) early in therapy was associated with better outcomes for the clients presenting with panic disorder. Lower therapist competence (and adherence) was found to be associated with dropout amongst the clients who presented with social anxiety. These findings suggest that both the therapists' competence and adherence to therapy protocols contributed to therapy outcome but at different phases in the treatment.

Despite some studies identifying a relationship between competence and therapy outcome, others have argued that this relationship is not as strong (Barber et al., 2007) or as consistent as expected (Shaw et al., 1999). In a study which examined the role of the therapeutic alliance in the relationship between therapist competence and outcome in brief psychodynamic psychotherapy, Despland et al. (2009) found no direct link between therapist competence and outcome. However, the findings suggested that alliance patterns were a moderator in the relationship between therapist competence and therapy outcome, which further emphasises the complexity of the relationship between competence and outcome.

The above literature focused on research conducted within adult psychotherapies, however there are similar inconsistencies within the child and adolescent psychotherapy literature. A systematic review which explored the literature on therapist adherence and competence process research in relation to CBT therapy outcome in children and young people (Rapley., & Loades, 2019) concluded that the literature was small and inconclusive. Amongst the studies reviewed, there were inconsistent findings with minimal to no effect sizes found for associations between therapist adherence, competence and outcome. Additionally, in a systematic review and meta-analysis assessing whether adherence and competence predicts youth outcomes, Collyer et al. (2019) found a small but statistically significant relationship between therapist adherence and therapy outcome. They highlighted that this association suggested that the maintenance of particular levels of therapist adherence was potentially important in the implementation of manualised interventions. However, they concluded that the small effect size that had been identified, indicated that therapy outcomes were likely to be strongly associated with factors other than therapist adherence. Specifically, they concluded that this small effect size was in line with the notion of common factors in psychotherapy, being that specific therapy model factors contribute relatively little to treatment outcome. Further to this, there was no significant relationship between therapist competence and therapy outcome.

Although the relationship between therapist adherence, therapist competence and client outcomes remain unclear in the psychotherapy literature overall, there is evidence of a link between treatment fidelity and substance use outcomes (see Table 1 for a summary of the literature). Martino et al. (2008) examined the adherence and competence of 35 therapists who delivered either MET or general drug counselling sessions to 461 clients who experienced substance use problems. Therapist's adherence and competence was measured

by the Independent Tape Rater Scale (ITRS; Ball et al., 2002), a measure of adherence and competence of the implementation of MI principles and techniques. The findings revealed significant, though modest (.13 to .34) associations between the therapist's fidelity to MI and client process variables and the outcome of therapy. Specifically, higher levels of therapist adherence and competence to MI was associated with an increase in client motivation, suggesting that therapist fidelity may enhance client motivation to change. In addition to this, higher levels of adherence and competence to MI skills in the MET condition was significantly related to negative drug urine screens. Similarly, McCambridge et al. (2012) examined whether differences in client's cannabis cessation three months following a brief MI intervention was attributable to the therapist's fidelity to the MI protocol. Audio recordings of MI sessions from 75 clients were assessed for therapist fidelity by two raters using the Motivational Interviewing Treatment Integrity (MITI) scale (Moyers et al., 2005). Findings indicated that after controlling for therapist effects, two particular aspects of therapist fidelity to MI were significantly predictive of cannabis cessation three months after the intervention, these were 'MI spirit' and 'complex reflections'. However, clients in the study were not randomly allocated to therapists, raising the possibility of differences between therapists in the participant's levels of receptivity to the MI intervention.

The above studies have highlighted the possibility of a relationship between therapist fidelity to MI based interventions and substance use outcomes. There have been similar findings within the literature investigating the relationship between therapist fidelity and substance use outcomes amongst adolescents. For example, Campos-Melady et al. (2017) investigated the adherence and competence ratings of 91 therapists and their relationship with the outcomes of 384 young people who had received a behavioural intervention, namely the Adolescent Community Reinforcement Approach (A-CRA). Ratings of therapists' adherence and

competence to the A-CRA model were measured by independent raters using the A-CRA procedures checklist. Multilevel models revealed that both between and within therapist competence was found to be significantly predictive of decreases in clients' days of substance use. This suggested that clients whose therapists delivered the intervention with greater competence compared to other therapists, and who delivered the intervention more competently than their own average, had larger reductions in substance use. Therapist adherence was not predictive of the client substance use outcomes in the full sample. However, when the analysis only included clients who had completed the 12-month follow-up, it was identified that the between therapist adherence score was predictive of a decrease in substance use. The researchers concluded that the study provided evidence for a relationship between the skill within which manualised behavioural interventions are implemented and substance use outcomes. Collectively, the above studies indicate that therapist fidelity may be positively related to client substance use outcomes.

There is increasing evidence for significant indirect links between treatment fidelity and variables that are associated with client outcomes in the addiction field (Campos-Melady et al., 2017). For instance, in a summary of substance use treatment fidelity, Brown and Lent (2008) reported that adherence and competence were associated with therapeutic alliance, client motivation and belief in the effectiveness of treatment. Furthermore, Barber et al. (2006) investigated the linear and curvilinear relations between therapist adherence, competence, therapy outcome and the interactions of these on the quality of the therapeutic alliance for clients who had received individual drug counselling (IDC) as part of the National Institute on Drug Abuse Collaborative Cocaine Treatment Study (Crits-Christoph et al., 1999). Independent raters who were considered to be experts in IDC, were selected to assess the adherence and competence of 12 therapists who delivered IDC to 95 clients.

Ratings of therapist fidelity were made using the Adherence Competence Scale for IDC for Cocaine Dependence (ACS-IDCCD; Barber et al., 1996). The analysis identified that higher therapist adherence to IDC was marginally associated with better outcomes as measured by the Beck Depression Inventory (BDI; Beck et al., 1991) but not the Addiction Severity Index (ASI; Fureman et al., 1990). There was support for a curvilinear relationship between adherence and outcome, in that low and high levels of adherence were associated with worse client outcomes, indicating that intermediate levels of therapist adherence were associated with the best outcomes. Interestingly, the analysis also highlighted that for some of the clients who had reported a strong therapeutic alliance, therapist adherence to the treatment was deemed to be essentially irrelevant to outcome, as those clients demonstrated improvements at the end of IDC. In contrast, when the therapeutic alliance was weaker, a moderate level of therapist adherence was associated with the best client outcomes. This suggests that therapeutic alliance may have complex moderating effects on the relationship between therapist adherence and treatment outcome.

Table 1: Summary of substance use literature examining the fidelity outcome relationship

Author and year	Aims	Clinical Population	Participants	Aspects of Fidelity	Fidelity measure	Findings
Barber et al. (2006)	To test hypotheses relating to linear and curvilinear relations among adherence, competence and outcome.	Adults with drug (cocaine) use problems	Clients (n= 95) Therapists (n= 12) who delivered IDC	Adherence Competence	ACS-IDCCD; Barber et al. (1996)	Higher adherence to IDC was marginally associated with better outcome as measured by the BDI (Beck et al.,1961) but not the ASI (Fureman et al., 1990). Support for a curvilinear relation between adherence and outcome. Low and high levels of adherence were associated with worse outcomes. Intermediate levels of adherence associated with the best outcomes. Therapist competence did not predict substance use outcomes.
Campos-Melady et al. (2017)	To examine ratings of adherence and competence and their relationship with outcome.	Adolescent substance use problems.	Clients (n=384) Therapists (n=91) trained to deliver A-CRA	Adherence Competence	The A-CRA Procedures Checklist.	Therapist competence predictive of decrease in days of substance use. Adherence not predictive of substance use outcome. Between therapist adherence predictive of decrease in substance use at 12-month follow-up.

Martino et al. (2008)	To examine adherence and competence of interventions associated with MI and general counselling.	Adults with substance use problems.	Clients (n= 461) Therapists (n= 35)	Adherence Competence	ITRS (Ball et al., 2002)	Higher levels of MI adherence and therapists' competence in MET condition was significantly related to negative urine screens during 4-week treatment phase.
McCambridge et al. (2012)	To examine whether differences in cannabis cessation 3 months after MI session was attributable to therapist fidelity.	Adolescent (16-19 years) frequent cannabis users.	Clients (n=75) Practitioners (n=4)	Adherence Competence	MITI scale (version 2)	Two aspects of enhanced fidelity (MI spirit and complex reflections) were predictive of cannabis cessation 3 months after MI session.

A-CRA = Adolescent Community Reinforcement Approach; **ACS-IDCCD** = Adherence Competence Scale for IDC for Cocaine Dependence; **ASI** = Addiction Severity Index; **BDI** = Beck Depression Inventory; **IDC** = Individual Drug Counselling; **ITRS** = Independent Tape Rater Scale; **MET** = Motivational Enhancement Therapy; **MI** = Motivational Interviewing; **MITI** = Motivational Interviewing Treatment Integrity scale

1.7 The complexities of researching the fidelity - outcome relationship

What is clearly highlighted from the inconsistencies in the literature is that the relationship between treatment fidelity and therapeutic outcome is complex. External variables which are not directly related to treatment fidelity can influence therapy outcomes, for example: severity of client symptomatology, client responsiveness, motivation and readiness to change, and/or the therapeutic alliance (Hogue & Dauber, 2013). Some authors argue that higher levels of symptom severity might require deviation from the treatment manual (e.g. Perepletchikova & Kazdin, 2005). Others argue that lower levels of adherence can be helpful for clients with low treatment motivation (Borrelli, 2011). For instance, in one study, higher levels of adherence were associated with poorer outcome in clients classified as having low motivation (Huppert et al., 2006). Multiple process variables (e.g. the therapeutic alliance), that are not always measured within studies, may contribute to the effect of adherence on outcome by masking or mediating significant associations (Webb et al., 2010). It is these confounding factors in addition to research design strategies that may explain the inconsistency of fidelity and outcome findings (Perepletchikova & Kazdin, 2005).

The importance of therapist adherence may vary across the course of therapy (Folke et al., 2017) and different therapeutic interventions that are applied in each session may require differing levels of adherence. For example, variability in the protocol during the earlier phases of therapy may be advantageous in motivating clients for engaging in the kind of therapy being delivered (e.g. Huppert et al., 2006). Furthermore, the role of therapist adherence in the outcomes of therapy may require interpretation based on the context in which specific elements of therapy models are delivered and understood. For example, in CBT for anxiety, exposure is a core element of the intervention that can be delivered in a

number of ways (e.g. through homework or face to face). If the therapist is not present during the exposure exercise than this part of preparation for the exercise may require more individual adaptation in advance compared to if the therapist guided the exercise face to face. In this instance, it is reasonable that the effect that adherence has on outcome may be influenced by how the treatment element is delivered. However, as Hauke et al. (2014) report, procedural variations on adherence are often not described in the literature.

There is the possibility that therapist competence moderates the relationship between adherence and outcome (Barber et al., 2007). For example, it may be that as therapist competence increases, adherence has a stronger or weaker relation with outcome. In this case, if competence is very high or very low then it might not matter how much the therapist is adhering to the treatment model. It is possible therefore that attempts to discover the relationship between adherence and outcome might fail if researchers do not take competence into account in research studies.

Inconsistencies in the literature investigating the association between competence and outcome may be due to statistical, methodological and/or conceptual factors. The relationship between competence and outcome may in fact be small and therefore difficult to detect. There is also the possibility that the effect of competence on outcome in many of the studies reviewed may have been small due to a restricted range of therapist competence. A large proportion of the literature looks at competence and outcome based on RCTs which tend to rely on experienced and/or well-trained therapists (e.g. Huppert et al., 2006; Trepka et al., 2004). This is likely to lead to low variability in competence and/or outcome (Brown et al., 2013). Like adherence, competence may have a different relationship to outcome at different stages of the therapeutic intervention (Barber et al., 2007). Examining the value of

measuring therapists' fidelity and client treatment outcome over time may enable researchers to understand the longevity, strengths and weaknesses of treatment fidelity (Prowse & Nagel, 2015). What has become apparent from completing a systematic search of the literature, is that most of the studies do not share the same method of analysing the relationship between competence and outcome. For example, Huppert et al. (2006) examined zero-order effects of competence on outcome, whilst Holder et al. (2018) utilised hierarchical linear modelling. In a number of cases, instruments used to measure adherence and competence were not psychometrically evaluated which raises questions around the reliability and validity of the ratings (Barber et al., 2005).

Finally, inconsistencies in the relationship between competence and outcome may also be due to conceptual issues. Adherence and competence are conceptually distinct. However, the degree to which they are separate across different measures is not consistent. Adherence may have a complex relationship with outcome (Barber et al., 2006) and failing to separate out adherence may obscure any relationship between competence and outcome (Barber et al., 2007). Furthermore, Barber et al. (2007) suggest that competence may not have a direct relationship with outcome but may moderate the effects of process variables like the therapeutic alliance or adherence which may be predictors of outcome in the presence of therapist competence.

1.8 Fidelity Measures in the Addiction Field

The addiction field was considered to be at the forefront of "advancing in addressing the issue of adherence and competence" in how therapies are delivered (Madson & Campbell, 2006, p. 67). The majority of fidelity measures that have been developed within the

addiction field were done so specifically for research purposes. These measures have been important in evaluating therapist adherence and/or therapist competence when implementing a particular treatment within a research trial. In the United Kingdom Alcohol Treatment Trial (UKATT), the UKATT Process Rating Scale (PRS) (Middleton et al., 2001) was developed for the purposes of the research trial. This measure was designed to assess both adherence and therapist competence when delivering MET and SBNT. Similarly, in the Addressing Drinking Among Patients: comparing Two Approaches (ADAPTA) trial (Watson et al., 2015) the ADAPTA Process Rating Scale (PRS) (Tober & Crosby, 2014) was designed to assess therapists' fidelity (adherence and competence) to an alcohol focused intervention and a health living intervention.

As previously discussed, establishing measurement tools which provide both reliable and valid data may enhance the assessment of competent practice by therapists in routine care (Falender & Shafranske, 2004). Such measurement tools would assist clinical supervisors to make empirically based evaluations of therapists under their supervision, as opposed to relying on clinical experience. Quality measurement of treatments may be particularly important within the substance use field where there is variability amongst the clinical and training experiences of treatment providers (Culbreth, 1999). However, although there have been fidelity measures developed for research purposes in the addiction field, there are a limited number of measures developed for both research and other purposes such as supervision, and self-reflective practice. For instance, the CBT Therapist Checklist (Carroll, 1997) was specifically designed for cocaine dependency and assesses therapist adherence in the delivery of CBT for this. Furthermore, the Contingency Management (CM) Clinician Rating Form (Petry & Stitzer, 2002) was designed for both research and training, and measures therapist adherence and competence to CM when treating substance use. There are

very few fidelity measures within the addiction field that are specifically for use in routine clinical practice and in enhancing therapist skill.

In addition to this, in routine clinical practice, therapists that deliver interventions for alcohol and drug use problems often use a range of therapies (Raistrick et al., 2006). In order to meet individual client needs, it is not unusual for therapists to tailor their therapeutic interventions (Norcross & Wampold, 2011) by drawing flexibly on techniques from different therapy models. Therefore, for a fidelity measure to have utility in routine clinical practice for substance use problems, a transtheoretical scale that evaluates therapist delivery of evidence-based interventions is required.

1.9 Brief Addiction Therapist Scale (BATS)

The BATS is a transtheoretical fidelity measure designed for use by clinicians practicing in drug and alcohol addiction services. As described above, in routine clinical practice, therapists typically use a range of psychological therapies to address clients' drug and alcohol problems (Riastrick et al., 2006); these can range from psychoeducation to intensive specialised interventions such as MET or CBT. A transtheoretical fidelity measure that can be used across a range of therapeutic models is particularly important within the context of the NHS where clinicians tend to be eclectic and flexible in their choice of treatment models in practice.

The BATS was developed to evaluate therapist delivery of evidence-based therapies widely used for addressing alcohol and drug use problems in routine clinical practice. As the BATS is transtheoretical and therefore does not focus on model specific interventions, the items on

the measure are grouped and loosely based on the transtheoretical stages of change model (DiClemente & Prochaska, 1998): i) items relevant to most sessions irrespective of a client's readiness to change, ii) items applicable for building motivation for change and iii) items appropriate for planning and maintaining change. Furthermore, considering the "common factors" previously discussed, the items on the BATS are consistent with the widely identified "common factors" in psychotherapy. For example, item two on the BATS, 'collaboration', refers to the extent to which the therapist attempted to work jointly with the client.

The BATS was developed for use in both supervision and training, with an aim to enhance therapist skills in delivering psychotherapy. Wampold and Imel (2015) propose that variance in therapeutic effectiveness is likely to reside in the therapist more than the specific treatment model being used in therapy. It seems sensible therefore that efforts should be directed towards developing therapists' skills. Indeed, a study investigating the characteristics of highly effective psychotherapists found that the amount of time that was spent on improving therapeutic skills predicted positive client outcomes (Chow et al., 2015). It is possible that fidelity measurement may be helpful in informing clinical supervision, and this is what the BATS was designed to do. For example, clinical supervisors can review a therapist's session and rate this using the BATS. This feedback can then be shared with the therapist and may consequently result in beneficial therapy outcomes for clients. Consequently, the aim of the current research study is to evaluate whether the BATS is associated with therapeutic outcome.

The BATS was developed by Crosby (2018) through several stages. Firstly, a literature review was conducted in order to identify pre-existing fidelity measures that evaluated the

delivery of psychological therapies used for treating alcohol and drug use problems. The literature review highlighted twenty-six fidelity measures in total, and it was these that formed the basis for generating the BATS items. This item pool was gradually refined, and consensus on the content of the BATS was generated amongst experts in the fields of addiction and psychotherapy, resulting in 12 items. The extensiveness of a therapist's behaviour for each of the BATS items is rated on a five-point Likert scale that ranges from 'not at all' to 'extensively'.

The BATS was subject to investigations examining its psychometric properties by Crosby (2018). More specifically, the BATS has undergone convergent validity and Inter Rater Reliability (IRR) analyses. The convergent validity of the BATS was examined in two ways. Firstly, it was examined by exploring the relationship between the BATS and the Working Alliance Inventory (WAI; Horvath & Greenberg, 1989), which is a well-validated measure of therapeutic alliance. Secondly, the total scores of the BATS were compared with the total scores from three process rating measures, which were the ADAPTA PRS (Tober & Crosby, 2014), the AESOPS Process Rating Scale (PRS) (Tober & Crosby, 2011), and the UKATT PRS (Middleton et al., 2001). There was evidence for convergent validity from both the WAI and two process measures (ADAPTA PRS and AESOPS PRS). Inter-rater reliability of the item scores were examined using the Interclass Correlation Coefficient (ICC) and weighted kappa. ICCs demonstrated good to excellent levels of reliability and weighted kappa coefficients indicated good to very good reliability between two raters.

1.10 Current study: primary and secondary research questions

Through using video and audio data acquired through the UKATT study, my thesis conducted a secondary analysis to examine if there was a relationship between the BATS ratings and therapy outcome. My study aimed to investigate whether the BATS scores could predict therapy outcome. It was hypothesised that there would be a positive relationship between therapist scores on the BATS and therapy outcome.

Data from clinical trials has increasingly been reanalysed to investigate estimates of the proportion of variability in outcomes that can be attributed to therapists delivering interventions (e.g. Crits-Christoph & Mintz, 1991; Huppert et al., 2001; Wampold & Brown, 2005). My secondary research question aimed to examine whether the therapists' BATS scores could explain the level of variation in outcomes.

Chapter Two: Method

2.1. Design

My project involved the secondary analysis of existing randomised controlled trial (RCT) data from the United Kingdom Alcohol Treatment Trial (UKATT Team, 2005), to determine the relationship between the BATS and the outcome of therapy. The UKATT study consisted of outcome data from three time points, baseline, three months and twelve months. My project utilised the UKATT outcome data from baseline and twelve months post intervention. Primary process data was generated through rating recorded therapy sessions from the UKATT trial using the BATS measure.

This current study adopted a positivist approach. Positivism believes that there is a "relationship between the world (objects, events, phenomena) and our perception, and understanding, of it" (Willig, 2013, p. 3). For positivists, knowledge of the world can be explored and measured from experience. This stance is dominant within the natural sciences, such as in health research, particularly through the use of quantitative methods.

2.2. Secondary Analysis Data

2.2.1. UKATT trial

The data for the secondary analysis came from UKATT (UKATT Team, 2005). This is the largest trial of treatments for alcohol problems that has been conducted within the United Kingdom (UK). UKATT was a multicentre RCT which aimed to identify whether

Motivational Enhancement Therapy (MET) was as effective as Social Behaviour and Network Therapy (SBNT). The trial also examined whether SBNT was as cost-effective as MET. MET was a manual-based therapy adapted from Motivational Interviewing (MI) (Miller & Rollnick, 2013), which aimed to reinforce a client's motivation and commitment to change their drinking behaviours (Miller et al., 1992). SBNT was an integrative approach which combined cognitive and behavioural techniques (Copello et al., 2002). This intervention aimed to help clients to build networks of individuals who would support positive change in their drinking behaviours (UKATT Research Team, 2001).

The UKATT trial recruited a total of 742 participants who attended specialist services for alcohol difficulties. Participants were randomly allocated to either the MET or SBNT intervention. The outcome data for the trial was collected at three different time points (baseline, three months and 12 months). Trial results revealed that both groups reported considerable reductions in alcohol consumption, dependence, difficulties, and better mental health related quality of life over the 12-month period. There was only one significant difference in outcome between the groups, with those in the SBNT arm of the trial demonstrating significantly better physical health at three months. The UKATT researchers concluded that the two interventions did not differ significantly in their effectiveness (UKATT Team, 2005).

The UKATT trial produced a wide range of data on the outcomes of the therapy sessions and the processes which occurred within them. This included a measure of change in client alcohol consumption, alcohol dependence and alcohol related problems over the 12 months. Therapy sessions were recorded on videotape (later transferred to DVD) and rated by UKATT researchers using the UKATT Process Rating Scale (UKATT PRS) (Middleton et

al., 2001), in order to assess the therapists' fidelity to the model. Four hundred and fifty-two sessions were rated as part of the trial, reflecting 27% of the total number of sessions (*n*=1664).

The UKATT data was selected for use within this current study due to it having a large number of accessible therapy session recordings. In addition to this, the trial included a range of therapists that delivered the interventions. The UKATT data used within this current study comprised i) recordings of MET therapy sessions across all original trial sites and ii) the primary outcome data collected in UKATT, the Form 90 (Miller, 1996).

2.2.2. UKATT trial participants

In this current study, I did not have any direct contact with the original UKATT participants. I used digital recordings of the therapy sessions which occurred as part of the original UKATT trial. The participants in this thesis research were a subsample of participants that had been allocated to the MET intervention in UKATT. This included the patients that received the MET intervention and the therapists that delivered it.

The participants that received the MET intervention were all over the age of sixteen, had alcohol abuse as their primary problem and would normally have received support from one of the UK treatment sites. The UKATT team had gained written consent from each participant for the storage of anonymised data and video recordings of their therapy sessions to be used in future research. The consent form from the original UKATT trial can be viewed in Appendix D. In UKATT participants were assigned with an identifying code. This code was also used in this research. The participants that received the MET intervention remained

anonymous in the current project as they could not be seen on the therapy recording, only their voice could be heard.

The trial therapists that delivered the MET intervention to participants were selected from the clinicians that were already employed by the participating treatment services across Birmingham, Cardiff and Leeds. Therapists were selected to participate in UKATT by submitting a curriculum vitae and a video recording of them demonstrating MI skills. All the trial therapists demonstrated evidence of having two years of practice in the addiction field, in addition to being deemed to have 'good therapeutic ability' and an ability to work with two or more clients simultaneously. Trial therapists attended a three-day standardised introduction to the trial and its procedures. They also received training in MET, to which they had been randomly assigned, which included role play and feedback. Following the training therapists were required to complete MET treatment with at least two clients before being assessed as competent in practicing in the trial.

2.2.3. Content of therapy recordings: Motivational Enhancement Therapy (MET)

MET is a manualised treatment based on MI principles and was adapted specifically for UKATT (Tober et al., 2002). In UKATT MET included three sessions. The first session included individualised feedback from pre-treatment client self-report questionnaires and results from their liver function tests. This feedback was often followed by a discussion around the client's perceived benefits of their current drinking behaviour, and problems associated with this, in order to elicit any of the client's concerns. The second MET session occurred one week after the first session and aimed to establish client goals and any anticipated barriers to achieving these goals. The third and final session occurred

approximately six weeks after the second session. During the final session, any specific instances of the client having abstained from drinking were reviewed alongside any continued problem drinking behaviours. The final MET session typically included an overall summary of progress with a view to elicit the client's optimism in the longer term. All sessions included the application of MI principles and strategies to enhance motivation in order to maintain positive change.

The MET therapy tapes were selected for my project for several reasons. Firstly, MET is a derivation of MI, an intervention strategy which is frequently used and viewed as useful for the treatment of a wide range of lifestyle problems (e.g. Vasilaki et al., 2006). Secondly, the MET intervention has also been identified as effective as more intensive treatments for people in alcohol specialist services (UKATT Team, 2005). The MET tapes were also chosen due to their short length, which made my project feasible. In UKATT participants attended three MET sessions, all of which lasted approximately 50 minutes. Selecting tapes that were of a relatively short length allowed me to rate more therapy tapes using the BATS, which was particularly important given the Doctorate in Clinical Psychology time frame.

2.3. Primary Data Derived from the BATS

My research project observed a subsample of therapy session tapes from a number of cases in the MET arm of UKATT. During the observation of the therapy sessions I rated the therapists using the BATS measure, this resulted in primary process data.

2.4. Sampling of cases

The MET therapy recordings from the UKATT trial were initially selected using a combination of random and purposeful sampling (See figure 1). I randomly selected thirty-five completed MET therapy cases from the UKATT database. 'Completed cases' refers to cases being selected on the basis that the participant had completed the MET intervention in UKATT i.e. they had attended all three MET sessions. I selected these by identifying all of the completed cases in the UKATT database. I then obtained a random selection of the completed cases by using a random number generator on SPSS version 23 (IBM Corp, 2015). However, out of the MET completed cases that were randomised, only fifteen completed case recordings were found in storage. This was potentially due to the physical DVDs not containing the correct or full participant identifiers. Consequently, I identified all of the completed cases that were in storage, which resulted in thirty-nine completed case recordings being found.

If one or both therapy recordings (MET session one and three) were of poor quality, which meant that these were unable to be viewed appropriately, then this case was excluded from the sample. Furthermore, if the case did not comply with ethical approval (e.g. whereby the client could be seen on the therapy recording) then the case was excluded. There were three cases excluded from the sample for these reasons. Two of the thirty-nine cases were excluded from the sample as the therapist on the video provided the client with incorrect consent information (that no one else would observe their therapy tape), this was in addition to the client being visible on the recordings. Another tape was excluded due to the therapy recording being too quiet for me to clearly hear what was being communicated during the session.

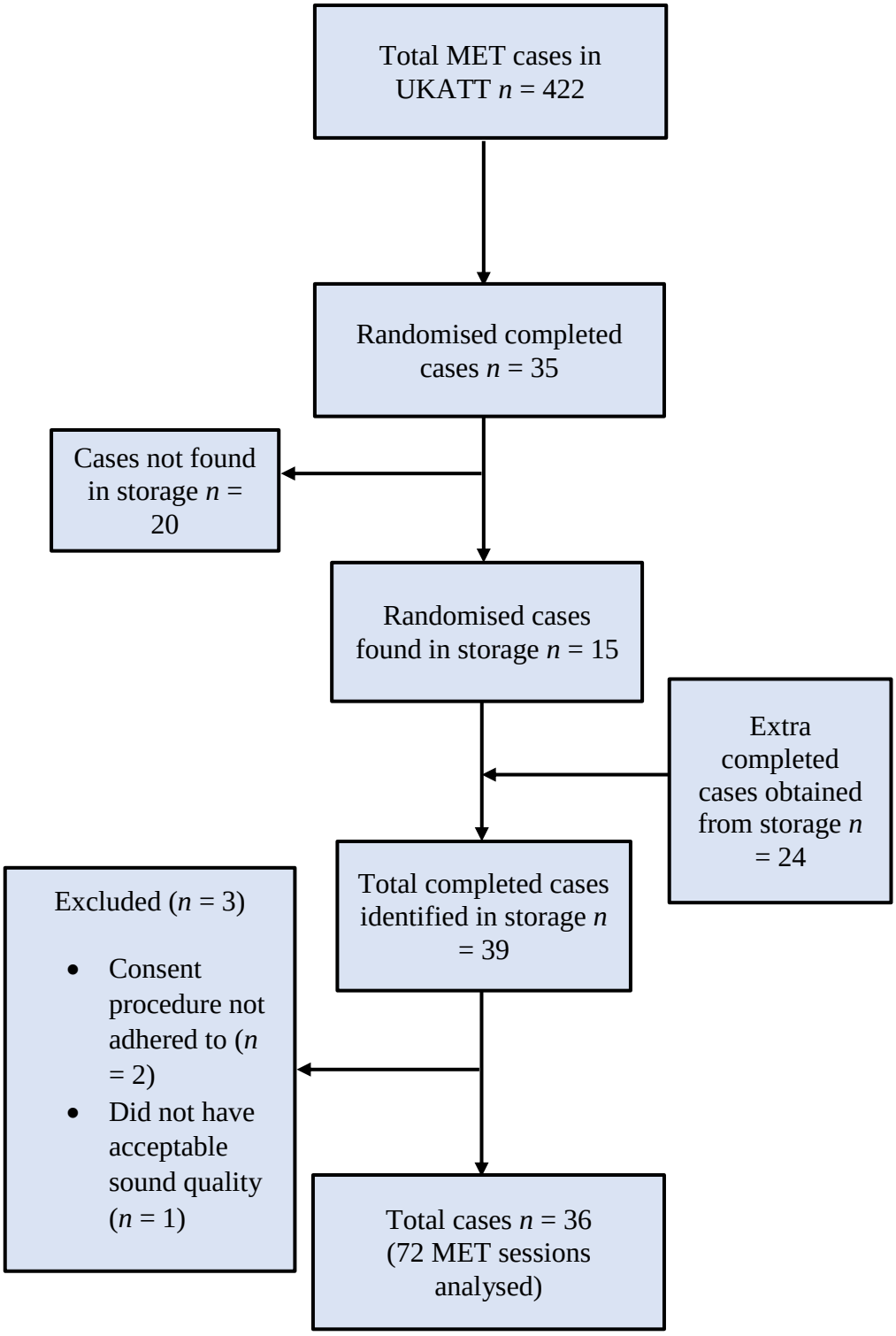
In summary, the MET recordings were excluded if:

1. They did not have acceptable sound quality
2. They did not have acceptable visual quality
3. Consent procedures for either UKATT or my project were not adhered to

In my study, an individual case was made up of two MET therapy session recordings as I observed two sessions from each participant that had completed the MET intervention. I selected therapy tapes from the first and final session of each case i.e. MET therapy session one and MET therapy session three. These will be referred to in my thesis as MET session one and MET session three. It was decided that rating two rather than three tapes would enable me to rate more individual cases. This would also potentially allow for the exploration of any differences in the behaviours that were being performed by therapists across the sessions and, if differences were identified, I could potentially explore whether the specific behaviours were related to outcome. I therefore observed and rated a total of seventy-two MET therapy session tapes from thirty-six different clients. The sample sizes were based on previous work that had explored psychometric properties of instruments (e.g. Torrey, 2012), in addition to what was feasible during the time frame.

Figure 1

Sampling of MET completed case recordings



2.4.1. Participants in my sample

I selected 36 clients from the UKATT database that had received and completed the MET intervention (i.e. they had attended all three MET sessions). All the participants would have normally received treatment for alcohol problems from the specialist treatment centres that took part in UKATT. The average age of the participants in my sample was 42 years. Seventy-two percent were male, and twenty-eight percent were female. This sample was similar to the overall sample of participants that had been randomised to the MET intervention in UKATT. The average age of the 422 participants that had been randomised to MET in UKATT was 42 years. Seventy-five percent were male, and twenty-five percent were female (UKATT Team, 2005).

2.4.2. Therapists in my sample

31 therapists (21 female and 10 male) delivered the MET intervention in UKATT. In my study, there were thirteen different therapists (8 Female and 5 Male) that were observed and rated using the BATs. The number of cases per therapist that I observed and rated ranged from once to eight times. Therapists were observed from all three trial sites in UKATT (Birmingham, Cardiff and Leeds). Birmingham was used four times, Cardiff three times, and Leeds six times. Seven of the therapists were reported in UKATT to have a 'Diploma' (three of which were classed as a 'Graduate Diploma'), three therapists were reported to have had a 'Degree', two had a 'Masters' and one therapist was reported to have a 'Post Graduate' level of education. It was not clear how some of these qualifications differed from one another. Therapists had, on average, 243 days of training. The therapist characteristics can be viewed in Table 2.

Table 2: Therapist Characteristics

Therapist	Number of cases rated	Sex	Trial site	Education	Training (in days)
1	2	Female	Birmingham	Diploma	114
2	5	Male	Birmingham	Degree	207
3	2	Female	Birmingham	Diploma	317
4	1	Female	Birmingham	Diploma	272
5	2	Male	Cardiff	Masters	405
6	1	Female	Cardiff	Post Grad	405
7	3	Female	Leeds	Graduate Diploma	333
8	3	Male	Leeds	Graduate Diploma	189
9	8	Female	Leeds	Degree	177
10	3	Male	Leeds	Graduate Diploma	233
11	2	Female	Leeds	Degree	64
12	3	Male	Leeds	Diploma	112
13	1	Female	Cardiff	Masters	333

2.5. Measures

This study used two measures: one administered within the original UKATT trial and one used by the researcher for the current study.

2.5.1 The measure selected by the researcher

Brief Addiction Therapist Scale (BATS)

The sample of MET therapy tapes that were selected were rated using the BATS, a tool for evaluating therapists' delivery of psychological therapies for alcohol and drug use (Crosby, 2018). The BATS was designed for use by clinicians in both supervision and training, with the aim to enhance therapist skill. This measure also has the potential to impact on therapist competence and therapy outcome.

The BATS is composed of twelve items which reflect the key features of therapies widely used in the addiction field such as MI. Initial psychometric work has been carried out on the BATS with preliminary analyses indicating that it has acceptable levels of inter-rater reliability, and support for face validity and convergent validity (Crosby, 2018).

I received training in how to score the BATS by the researcher (HC) who developed the measure. There was a calibration of ratings training phase whereby we observed therapy tapes from UKATT together. These included tapes from both the SBNT and MET arm of the trial. The rated tapes were then discussed whilst referencing the item definitions for the purposes of calibration (Watson et al., 2013). Four therapy tapes were rated in this way. It

was not until HC was satisfied that our scoring of the BATS was consistent, that I was signed off as competent in independently rating the therapy tapes. Inter-rater reliability (IRR) for BATS ratings was also established during this study and will be discussed further in this thesis.

2.5.2 The measure selected from the UKATT trial data

Form 90 (Miller, 1996)

The Form 90 measure allows the calculation of drinks per drinking day and percentage of days abstinent. The Form 90 was selected for the proposed study due to it being used as the primary outcome measure of alcohol consumption in UKATT. During UKATT, the Form 90 data was collected at baseline, three and twelve months after entry to the trial. I requested access to the UKATT Form 90 outcome data from the University of York who are custodians of the data, access to which was granted.

The Form 90 has been subjected to two test-retest studies that revealed good to excellent reliability for all key summary measures of alcohol consumption and psychosocial functioning, whilst most frequently used illicit drugs demonstrated moderate reliability. The Form 90 is considered to be a reliable measure for alcohol treatment assessment research, particularly when those conducting the Form 90 interview had received thorough training and supervision in its use (Tonigan et al., 1997).

2.6. Ethical Approval

Ethical approval for this project was obtained from the University of Leeds (Reference MREC 18-077) on 9th May 2019 (Appendix E), with an amendment to ethics approved on 14th October 2019 (Appendix F).

There were two permission processes in order to gain access to the UKATT data for the purposes of secondary analysis. Firstly, the MET therapy tapes from UKATT reside in secure storage in the Division of Psychological Medicine at the University of Leeds in DVD format. The original NHS trial sites in the UKATT study no longer have any involvement or interest in the UKATT project or the study materials. I was granted access to the MET therapy tapes by one of the UKATT Principal Investigators (PI) who is an honorary member of staff at the University of Leeds.

Secondly, the UKATT trial primary outcome data from the Form 90 and demographic information for both participants and the therapists is held in a data repository at the University of York. I applied for access to this data which was accepted and granted by the University of York (Appendix G). This data was password protected and stored on my University of Leeds m: drive, which was accessed through the University of Leeds Desktop Anywhere.

2.6.1 Confidentiality

All the participant data from the UKATT trial was anonymised as each participant was assigned with a client identifying code by UKATT researchers. Therefore, informed consent for my study was unable to be obtained from the UKATT participants. The original ethics application form for UKATT allowed for secondary studies and analyses of the data. All participants in UKATT provided written consent for this by signing the UKATT consent form (Appendix D). Indeed, several secondary analyses have been conducted on this existing data to date.

The same participant identifying code that was allocated to participants in UKATT was utilised in my study. The participants that received the MET intervention in UKATT cannot be seen on the therapy tape recordings. Only the therapist delivering the MET intervention can be seen on the recording of the therapy sessions. Participants receiving treatment are audible on the tape but remained anonymous to me. There were two instances whereby the participant receiving treatment was visible at the start of the tape. Consequently, these tapes were not used in my research project.

2.6.2. Data Storage

The data used in my project was sensitive, and therefore confidentiality was maintained, and the data kept secure. All viewing and analyses of the UKATT participant therapy recordings and the data originating from these was completed in a private, confidential space. This was most frequently at the site where the recordings were securely stored at the University of Leeds. However, IRR ratings were occasionally completed in a private and confidential space

by my thesis supervisor (HC) either at their home address or place of work (another university setting). This was approved in the ethical amendment (Appendix F). These particular tapes for IRR were transferred onto an encrypted memory stick which was password protected and contained no identifiers.

The BATS measure was completed during the viewing of the MET therapy session recordings by myself, and my supervisor. The completed BATS form did not contain any identifiable information but corresponded to the relevant tape, which contained the participant and therapist identifying codes that were used in UKATT. The anonymised completed BATS forms were stored in the Doctorate in Clinical Psychology office at the University of Leeds. A database was also created with the abstracted information, which I accessed through Desktop Anywhere.

2.7. Data collection

2.7.1. Procedure

Once the MET therapy recordings to be used in my study were selected, the BATS was used to rate the therapists' approach in each session (MET session one and MET session three). I listened to each recording (n=72) and scored the extent to which the therapists demonstrated each of the BATS item behaviours. Individual notes were made during the observation of each session in order to support the scores given to the BATS items. The items were scored at the end of the therapy session recording. Note taking was particularly useful for the sessions, which were longer in duration (i.e. those which exceeded fifty minutes, and/or those which had been selected for double rating).

I kept a record of the duration of the therapy session, the therapist, the participant and the site identifier, the date of the rating and the participant's goal when this was identified. Ratings took place between August 2019 and February 2020.

2.7.2. Reliability procedure

Twenty of the seventy-two therapy tapes were randomly selected for double rating. This reflected 28% of the total number of sessions. However, due to time constraints, ten of the twenty recordings were double rated. This reflected 14% of the total number of sessions.

The MET therapy recordings were rated independently by HC and me (KT). The rated recordings were then discussed with reference to the BATS item definitions for the purpose of calibration. Double rating took place at regular intervals between August 2019 and February 2020. Six tapes were rated during a meeting with the thesis supervisor and four were rated when at different sites. Session scores, duration and the date that the therapy recording was rated were entered into an excel spreadsheet.

IRR analyses tested the consistency of measurement on the BATS between the two different raters, KT and HC. IRR is the most commonly reported test of reliability and refers to the consistency of measurement between two or more raters (Trevethan, 2017).

2.8 Statistical analysis

Multilevel regression analyses of the primary process data (BATS ratings) and secondary data from UKATT (Drinks per Drinking Day and Abstinence data) were used to examine whether there was a relationship between the BATS and therapy outcome. These analyses were also utilised to investigate whether the BATS scores could explain the level of variation between therapists.

The BATS scores across each individual therapist's sessions were averaged to yield a total BATS score per therapist for MET session one and MET session three (Table 3). The mean BATS score for MET session one (BATS score 1) ranged from 14 to 25. The mean BATS score for MET session three (BATS score 2) ranged from 16.5 to 24.33.

Table 3: Therapist average BATS scores MET session one and three

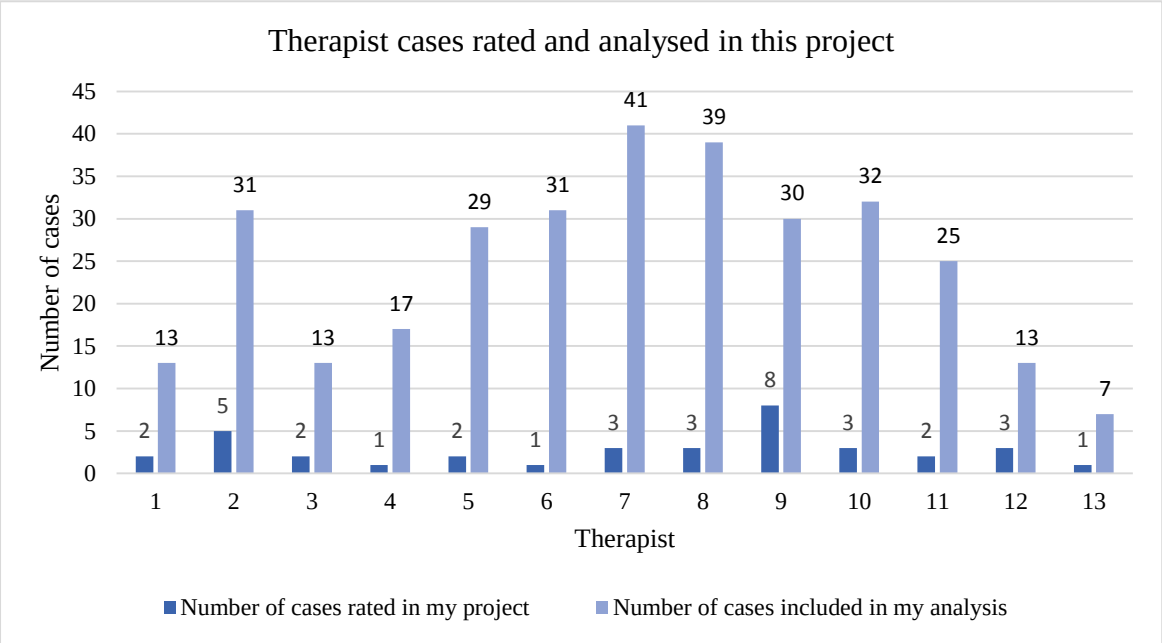
Therapist	Average BATS score MET session 1 (BATS score 1)	Average BATS score MET session 3 (BATS score 2)
1	20.50	22.50
2	22.40	22.40
3	17.50	16.50
4	21.00	20.00
5	16.50	24.00
6	25.00	18.00
7	19.33	20.00
8	18.00	22.33
9	22.63	24.12
10	15.66	19.00
11	20.50	23.50
12	20.66	24.33
13	14.00	20.00

The individual therapist's mean BATS score for MET session one and MET session three was then assigned to each participant that they had delivered the MET intervention to during the UKATT trial. This enabled me to investigate associations between their BATS data and the outcome data for all of their clients. The total number of participants that the thirteen therapists delivered the MET intervention to during UKATT was 321 participants. The number of participants that each of the selected thirteen therapists delivered the MET

intervention to in UKATT ranged from seven participants to forty-one, with therapist seven seeing the most clients and therapist thirteen seeing the least (Figure 2).

Figure 2

Therapist cases rated and analysed in this project



2.8.1 Multilevel Analysis

The data was analysed using a Random Intercept Model using STATA version 15 (StataCorp, 2017). This enabled me to investigate whether the therapists' BATS ratings from the MET session recordings could predict therapy outcome as measured by the Form 90 (Miller, 1996) outcome data from UKATT. The models provided the between and within therapist variance estimates necessary to calculate the Interclass Correlation Coefficients (ICC). ICCs are calculated as the ratio of variability between therapists to the total variability, which includes both therapist level and patient level error variability (Streiner et al., 2015).

This allowed for the examination of whether the correlation of participant outcomes within therapists was explained by the BATS scores for MET sessions one (i.e. BATS score one), MET session three (i.e. BATS score two) and a combination of these two BATS scores. The therapists in this study were considered to be a random factor. This was so that conclusions could be made about therapists in general (see Wampold & Brown, 2005). Whilst the therapists' mean BATS scores (for MET session one and three) were classed as a fixed factor. Other fixed effects were, the participants age, gender and their adherence (whether this was DDD or abstinence) as these were all included in the model.

The analysis included four multilevel models. These were as follows:

Model one: This performed a multilevel mixed-effects linear regression which investigated the relationship between the BATS scores and the participants Drinks per Drinking Day (DDD) data at 12 months. On inspection, the DDD data at 12 months was not normally distributed. Therefore, as this was classed as 'count' data, a log transformation was performed which transformed the data into a log of the raw values. The model also enabled me to examine the therapist and patient level variance. Through the calculation of the ICC, the proportion of the variance attributable to the therapists was investigated. This ICC was examined in a series of ways: i) not explaining the therapist level variance at all, i.e. not adjusting for the BATS scores, ii) explaining the therapist variance by adjusting for the therapists' BATS score one, iii) explaining the therapist variance by adjusting for the therapists' BATS score two, and iv) explaining the therapist variance by adjusting for a combination of BATS scores one and two.

Model two: In model two a multilevel mixed-effects linear regression was performed on the DDD data at 12 months when adjusting for baseline. The DDD data at 12 months and at baseline were log transformed due to the data not being normally distributed. The model enabled me to see whether the therapist BATS scores could predict the patient's outcome when adjusting for their baseline scores. The model also enabled me to examine the therapist level and patient level variance through the calculation of the ICC, when adjusting for some of the patient level variance (baseline data). The ICC for the therapist level variance was examined in four ways: i) not explaining the therapist level variance through the BATS scores, but accounting for some of the patient level error variance due to the inclusion of baseline data, ii) adjusting for the therapists' BATS score one, iii) adjusting for therapists' BATS score two, and iv) explaining the therapist variance by a combination of the BATS scores one and two.

Model three: In model three a multilevel mixed-effects logistic regression was conducted. This model aimed to investigate the relationship between the BATS and therapy outcome through the participant abstinence data. The participant's abstinence data was transformed into binary data whereby percentage abstinence less than 50% was coded as a 0 and percentage abstinence more than or equal to 50% was coded as a one. Due to the abstinence data being transformed into binary data, the multilevel mixed-effects logistic regression was selected as this is used to model binary outcome variables. The model enabled me to examine whether the BATS score could predict therapy outcome as represented by the abstinence data. It also enabled me to calculate the ICC in order to examine the therapist level variance. The ICC was investigated in four different ways: i) not explaining the therapist variance at all through accounting for the BATS scores, ii) explaining it by accounting for therapists' BATS

score one, iii) explaining it by adjusting for therapists' BATS score two, and iv) explaining the therapist variance by adjusting for a combination of the BATS scores one and two.

Model four: A multilevel mixed-effects logistic regression was performed on the abstinence data at 12 months and at baseline. The abstinence data at both baseline and 12 months were transformed into binary data whereby percentage abstinence less than 50% was coded as a 0 and percentage abstinence more than or equal to 50% was coded as a one. This model enabled the examination of whether the BATS scores could predict the patients' outcome (as represented by the abstinence data) when adjusting for their abstinence baseline scores. The model also enabled me to investigate the therapist level variance through the calculation of the ICC when adjusting for some of the patient level variance (baseline data). The ICC was examined in four ways: i) not explaining the therapist variance through the inclusion of BATS scores, but explaining some of the patient level error variance by adjusting for the patient abstinence baseline data, ii) adjusting for therapists' BATS score one, iii) adjusting for therapists' BATS score two, and iv) explaining the therapist variance by adjusting for a combination of the BATS scores one and two.

2.8.2 Inter-rater reliability analysis

Ten of the twenty recordings selected for IRR were double rated by my supervisor (HC), this was 14% of the total number of sessions. IRR of the item scores was examined by using the ICC, a test of agreement for continuous variables (Liu et al., 2016). ICCs are frequently used in the validation of fidelity measures (e.g. Watson et al., 2013; Carroll et al., 2000). To perform the IRR analyses, the ICC using a two-way mixed-effects model was utilised (Shrout & Fleiss, 1979). This model was selected because the MET therapy recordings were rated by

the same two raters (KT & HC) who were not randomly selected. There was also an effect of the therapist that delivered the MET session on the recording and was classed as a random effect, in addition to an effect of the raters (HC & HT) that were considered a fixed effect.

ICC estimates and their 95% Confidence Intervals (CI) were computed using SPSS (IBM Corp., 2013) and were interpreted by using the guidelines, which have been provided by Cicchetty (1994). The agreement guidelines for ICC are provided in Table 4.

Table 4: Agreement with ICC (Cicchetty, 1994)

ICC	Strength of agreement
<0.40	Poor
0.40 to 0.59	Fair
0.60 to 0.74	Good
0.75 to 1.00	Excellent

ICC = Inter class correlation coefficient; <=less than.

Chapter Three: Results

A secondary analysis of data from the United Kingdom Alcohol Treatment Trial (UKATT) (UKATT Team, 2005) was carried out. UKATT featured a total of 52 therapists who delivered one of two interventions (Motivational Enhancement Therapy, MET or Social Behavioural Network Therapy, SBNT) to 742 clients that attended specialist services for alcohol difficulties. The clients' outcome of therapy was measured in terms of Drinks per Drinking Day (DDD) (measured in units of alcohol) and percentage days abstinence.

In my study, 72 MET sessions (36 from MET session one and 36 from MET session three) delivered by thirteen therapists in UKATT were rated using the Brief Addiction Therapist Scale (BATS). BATS scores from MET session one and three for each therapist were averaged and assigned to all clients (a total of 321 clients) that each of the thirteen therapists saw in UKATT. The characteristics of the clients from the sessions I observed and the thirteen therapists in this study are provided in the method section of this thesis.

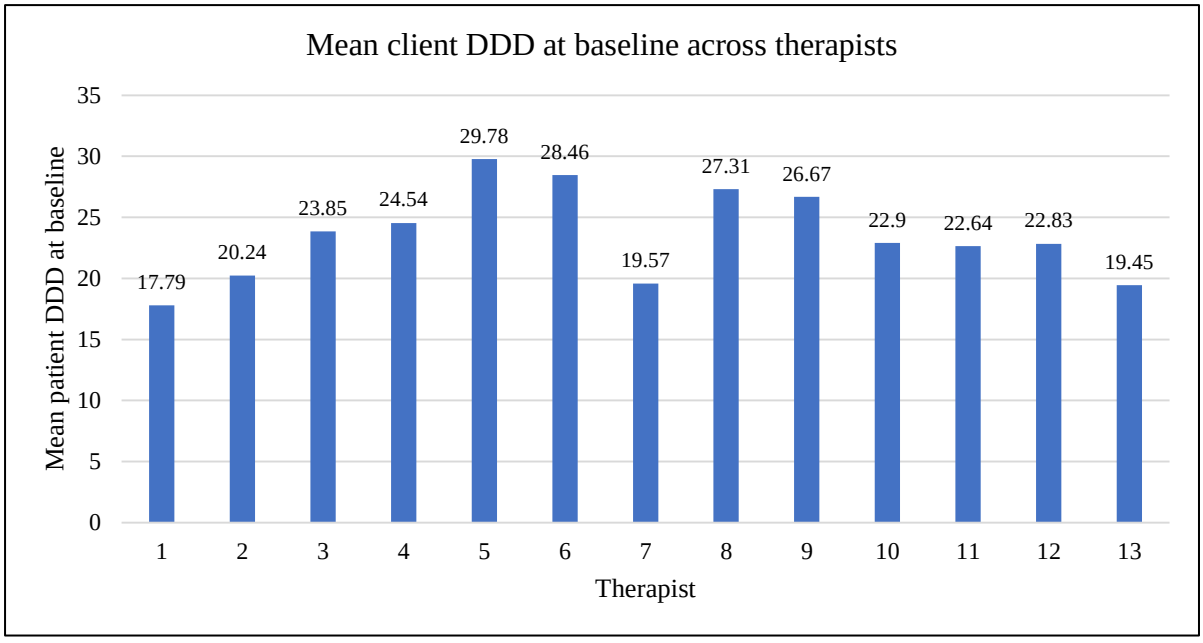
The following chapter reports information about the patient characteristics across the thirteen therapists, the BATS scores, and results from the Inter Rater Reliability (IRR) tests. This chapter also reports the findings of the multilevel models used to determine the relationship between the BATS scores and 1) patient outcome and 2) therapist variance in outcomes.

3.1 Patient characteristics across therapists

The thirteen therapists selected in my study saw a total of 321 clients in UKATT between them. The mean age of clients across the thirteen therapists ranged from 37.75 years to 44.51, and all of the therapists saw more male clients compared to female clients. The mean DDD data (measured in units) for clients at baseline across therapists ranged from 17.79 to 29.78, with therapist one seeing the clients with the lowest amount of DDD at baseline and therapist five seeing the most (see Figure 3). More detailed information regarding patient characteristics such as employment status, marital status, and education, across the thirteen therapists can be viewed in Appendix H.

Figure 3

Mean client Drinks per Drinking Day (DDD) at baseline across therapists



3.2 Sessions rated using the BATS measure

Descriptive statistics of the duration of the recordings for MET session one and three analysed in my study are given in Table 5. On average the first MET sessions were longer in duration when compared to the final MET sessions. The mean duration for MET session one, based on 36 recordings was 54.8 minutes (range 37 to 70 minutes). The mean duration of the final MET session, (MET session three) based on 36 recordings, was 39.1 minutes (range 14 to 55 minutes).

Table 5: Summary of the length of sessions for MET session one (*n* = 36) and three (*n* = 36)

	Duration of session (minutes)			
	Mean	**SD**	**Minimum**	**Maximum**
MET 1	54.8	8.2	37	70
MET 3	39.1	12.3	14	55

3.3 Item scores on the BATS

A summary of the item scores on the BATS across MET sessions one and three are presented in Table 6. Items one (session aims), two (working together) and three (conveying empathy) had relatively high scores across both MET sessions one and three. Items six (planning tasks), seven (reviewing tasks), nine (inconsistencies) and twelve (social network) had low scores across both MET sessions.

Table 6: Summary of BATS item scores for MET sessions one (*n* = 36) and three (*n* = 36)

Item reference	Therapy session	
	MET 1	MET 3
	Median Scores*, (IQR), Range	Median Scores*, (IQR), Range
1. Session aims	4 (1), 2-4	3 (1), 1-4
2. Working together	4 (1), 3-4	3.5 (1), 1-4
3. Convey empathy	4 (1), 2-4	3 (1), 1-4
4. Strengths	1 (1), 0-3	1 (2), 0-3
5. Complex reflections	2 (1), 1-4	2 (2), 1-4
6. Planning tasks	0 (0.5), 0-3	0 (0), 0
7. Reviewing tasks	0 (0), 0	0 (0), 0-1
8. Treatment goals	3 (1), 1-4	3 (1), 1-4
9. Inconsistencies	0 (0), 0-1	0 (0), 0-2
10. Talking about change	2 (3), 0-4	3 (2), 0-4
11. Change planning	0 (1), 0-3	2 (1), 0-3
12. Social network	0 (1), 0-1	0 (1), 0-3

*Scores made on a 5-point scale: 0=not at all, 1=a little, 2=somewhat, 3=a good deal, 4=extensively. IQR = Interquartile range.

3.4 Inter Rater Reliability (IRR)

A sample of the MET recordings were rated by two raters (KT & HC) (*n* = 10). This was equivalent to five MET session one recordings and five MET two recordings from six different participants and five different therapists. IRR was examined using the Intraclass

Correlation Coefficient (ICC) two-way mixed-effects model (Shrout & Fleiss, 1979). ICCs for individual items ranged from 0.86 to 1.00 indicating an excellent level of agreement between both raters (Table 7). The 95% CIs indicated some uncertainty around the magnitude of effect for item 11 ('change planning') as the CIs for this item ranged from fair to excellent (0.44 to 0.96). The ICC for agreement of total BATS scores was excellent (ICC = 0.99).

Table 7: Intraclass correlation coefficient for items on BATs and total BATS scores

Item Reference	ICC (95%) CI
1. Session aims	0.95 (0.79 to 0.99)
2. Working together	0.97 (0.89 to 0.99)
3. Convey empathy	0.97 (0.89 to 0.99)
4. Client strengths	0.98 (0.92 to 0.99)
5. Complex reflections	0.93 (0.70 to 0.98)
6. Planning tasks	1.00 (1.00 to 1.00)
7. Reviewing tasks	1.00 (1.00 to 1.00)
8. Treatment goals	0.96 (0.83 to 0.99)
9. Inconsistencies	1.00 (1.00 to 1.00)
10. Talking about change	0.97 (0.87 to 0.99)
11. Change Planning	0.86 (0.44 to 0.97)
12. Social Network	1.00 (1.00 to 1.00)
Total BATS Score	0.99 (0.96 to 1.00)

ICC = Intraclass correlation coefficient; CI = Confidence Interval

3.5. Multilevel analyses

Multilevel analyses of the primary process data (BATS ratings) and secondary data from UKATT (DDD and Abstinence data) were used to examine whether there was a relationship between the BATS and therapy outcome. Multilevel analyses were also utilised to explore whether the BATS scores could explain the level of variation between therapists. The analysis consisted of four models.

3.5.1 Model one: Multilevel mixed-effects linear regression

Model one performed a multilevel mixed-effects linear regression, which investigated the relationship between the BATS scores and the patient DDD outcome data at 12 months. The model identified that there was DDD outcome data at 12 months for 234 cases (72.89%) out of the total 321 cases that the thirteen therapists saw. Out of these cases patients per therapists ranged from four to thirty-four, indicating a lot of variability.

The model identified that the BATS scores did not predict therapy outcome as $p > 0.05$ for each of the fixed effects (Table 8).

The model also enabled the examination of the therapist level variance and the patient level error variance for the DDD outcome data at 12 months. The analysis yielded the estimates necessary to determine the proportion of variability that was due to the therapists. In the unadjusted model (i.e. not accounting for the therapist BATS scores), the estimate of therapist variance denoted by $\hat{\sigma}^2_{\text{ther}}$ was equal to 0.012 (*SE* = .016). The estimate of the patient error level variance denoted by $\hat{\sigma}^2_{\varepsilon}$ was equal to 0.57 (*SE* = .054).

The proportion of variance that was attributable to the therapists was given by the calculation of the ICC, defined in the following way (Snijders & Bosker, 1999; Wampold & Serlin, 2000):

$$\hat{\rho}_1 = \frac{\hat{\sigma}^2_{\text{ther}}}{\hat{\sigma}^2_{\text{ther}} + \hat{\sigma}^2_{\varepsilon}} = \frac{0.012}{0.012 + 0.57} = 0.021$$

This coefficient is the estimate of the population proportion of variance due to therapists divided by the total variance (both therapist variance and patient and error variance). The proportion of therapist variance for the unadjusted model was small at 0.021.

When the therapist BATS scores were added into the model, the proportion of variance due to the therapist ranged from 0.013 to 0.018 (Table 8). When the therapists' mean BATS scores from session one was included in the model as fixed effects, the ICC for the therapist variance was 0.016, reflecting a 23.81% decrease from the unadjusted model therapist variance. When adjusting for BATS score two in the model, the ICC for the therapist variance was 0.018, which was a 14.29% decrease from the ICC for the unadjusted model. When a combination of both BATS scores (one and two) were accounted for, the therapist variance reduced to its smallest amount at 0.013, this represented a 38.1% decrease from the unadjusted therapist variance. This suggested that the BATS scores might explain a portion of the therapist variance.

Table 8: Model of fixed effects and their relationship with therapy outcome for DDD at 12 months

Fixed effects	Estimate	*SE*	*z*	*p*	95 CI	ICC
BATS 1	.015	.021	0.73	0.47	-.025- .055	0.016
BATS 2	.012	.025	0.49	0.62	-.036 - .061	0.018
Combined BATS 1 & 2						
BATS 1	.014	.020	0.71	0.48	-.026 - .054	0.013
BATS 2	.011	.024	0.47	0.64	-.036 - .058	

CI = Confidence Interval; DDD = Drinks per Drinking Day; ICC= Intraclass Correlation Coefficient; SE = Standard Error

3.5.2 Model two: Multilevel mixed-effects linear regression

Model two performed a multilevel mixed-effects linear regression. This investigated the relationship between the BATS scores and the patient DDD data at 12 months whilst including an adjustment for the patients' DDD baseline data. When adjusting for the DDD data at baseline, the BATS scores did not predict therapy outcome as $p > 0.05$ for each of the fixed effects (Table 9).

The model enabled the examination of the variance relating to the therapist when adjusting for the patient DDD data at baseline. This allowed us to explain some of the patient level error variance by accounting for the patient baseline values. In doing so, the patient level error variance decreased by 24.56%, from 0.57 in the unadjusted model (model one) to 0.43 in the current model.

The proportion of therapist variance for this model when only adjusting for the patient level error variance, i.e. when excluding BATS scores, was small (ICC = 0.011). When the therapists' mean BATS scores were accounted for, the proportion of variance that was due to the therapist ranged from 0.001 to 0.008. When the therapists' BATS score one was included in the model, the ICC for the therapist variance fell to 0.004. This represented a 63.64% decrease from the ICC when the BATS scores were not accounted for. When only BATS score two was included in the model the ICC for the therapist variance was at 0.008 which reflected a 27.27% decrease. When a combination of both BATS scores (BATS one and two) were accounted for the therapist variance fell to its smallest amount at 0.001. This represented a 90.91% decrease from the ICC provided by the model when only adjusting for the patient level error variance.

Table 9: Model of fixed effects and their relationship with therapy DDD outcomes at 12 months when adjusting for DDD at baseline

Fixed effects	**Estimate**	***SE***	***z***	***p***	**95 CI**	**ICC**
BATS 1	.012	.016	0.75	0.45	-.020 - .044	0.004
BATS 2	.011	.020	0.53	0.59	-.028 - .050	0.008
Combined BATS 1 & 2						
BATS 1	.012	.016	0.74	0.46	-.019 - .043	0.001
BATS 2	.010	.019	0.53	0.60	-.027 - .047	

CI = Confidence Interval; DDD = Drinks per Drinking Day; ICC = Intraclass Correlation Coefficient; SE = Standard Error

3.5.3 Model three: Multilevel mixed-effects logistic regression

Model three performed a multilevel mixed-effects logistic regression, which investigated the relationship between the BATS scores and the patient percentage days abstinence data at 12 months. The model identified that there was abstinence data at 12 months for 272 cases (84.73%) out of the total 321 cases that the thirteen therapists saw. Out of these cases, patients per therapist ranged from six to thirty-eight indicating a lot of variability.

The model revealed that the BATS scores did not predict therapy outcome as $p > 0.05$ for each of the fixed effects (Table 10).

The model examined the variance components related to the therapists and the patient level error variance for the patient abstinence data at 12 months. The model yielded the estimates necessary to determine the proportion of variability due to the therapists. The proportion of variance due to the therapists was provided by the calculation of the ICC using a latent variable approach. The latent variable approach considers the observed binary response to represent a threshold continuous latent variable where we observe 0 below the threshold and 1 above the threshold. In other words, it assumes that every patient had a certain propensity for achieving abstinence but only persons whose propensity crosses a certain threshold did this. In this model we have an underlying logistic distribution for this latent variable. Logistic distribution has a variance of $\pi^2 / 3 = 3.29$**.** We can take this as level 1 variance, so that the level 1 (patient variance) and level 2 variance (therapist variance denoted by T'^2) are on the same scale. On this basis the ICC is then given by the following formula (Merlo et al., 2006):

$$\frac{T^2}{T^2 + \frac{\pi^2}{3}}$$

Table 10: Model of fixed effects and their relationship with abstinence outcomes at 12 months

Fixed effects	Estimate	*SE*	*z*	*p*	95 CI	ICC
BATS 1	-.008	.043	-0.19	0.85	-.094 - .078	1.38602e-35
BATS 2	.051	.053	0.93	0.35	-.054 - .153	4.86322e-35
Combined BATS 1 & 2						
BATS 1	-.012	.044	-0.27	0.79	-.098 - .074	4.13374e-37
BATS 2	.051	.053	0.95	0.34	-.053 - .155	

CI = Confidence Interval; ICC = Intraclass Correlation Coefficient; SE = Standard Error

This proportion of the therapist variance for the unadjusted model as provided by the ICC was very small at 4.40729e-33. When the BATS scores were added into the model the proportion of variance due to the therapist ranged from 4.13374e-37 to 4.86322e-35 (Table 10). When the therapists' BATS score one was included in the model, the ICC for the therapist variance was 1.38602e-35. This represented a 99.69% decrease from the unadjusted model. When only accounting for BATS score two in the model, the ICC for the therapist variance was 4.86322e-35, which was a 98.90% decrease. When a combination of both BATS scores (BATS one and two) were accounted for the therapist variance was at its smallest amount at 4.13374e-37. This reflected a 99.99% decrease from the ICC provided by the model when not adjusting for the BATS scores. Although the therapist variance in this model was much smaller than in models one and two, the variance followed a similar pattern

to the therapist variance when I examined this in relation to the DDD outcome data in models one and two.

3.5.4 Model four: Multilevel mixed-effects logistic regression

Model four performed a multilevel mixed-effects logistic regression to investigate the relationship between the therapist BATS scores and the patients' percentage day's abstinence data at 12 months when adjusting for their percentage day's abstinence data at baseline. The percentage day's abstinence data at baseline was omitted by the model in STATA due to collinearity. Therefore, I was unable to assess whether the BATS scores predicted outcome when adjusting for baseline data. Furthermore, I was unable to examine the therapist variance when accounting for some of the patient level error variance by adjusting for their baseline values. Consequently, the results for model four are not presented.

Chapter Four: Discussion

This chapter summarises the key findings of my thesis project within the context of the wider literature on fidelity measures and their relationship with treatment outcomes. The strengths and limitations of my project and how these issues may have influenced the research findings are discussed. The clinical implications of my project, including its limitations and potential areas for future research are discussed. The chapter ends with a summary of conclusions.

4.1. Summary of results

Multilevel analyses of primary process data from the Brief Addiction Therapist Scale (BATS) scores, and secondary data from United Kingdom Alcohol Treatment Trial (UKATT) (UKATT Research Team, 2005), were used to test whether there was a relationship between BATS ratings and therapy outcome. More specifically, I investigated whether the therapists' BATS scores could predict therapy outcome. I hypothesised a positive relationship between the BATS scores and outcome. Though the results of the analysis showed that the BATS scores did not predict therapy outcome, analyses relating to secondary research questions revealed a more complex picture, showing that BATS scores might explain some of the variance in outcomes.

4.2. BATS scores in the selected therapy sessions

Across both Motivational Enhancement Therapy (MET) sessions one and three (i.e. first session and final session respectively), the BATS scores were rated the highest on items one,

two and three. These items on the BATS measure relate to; i) the therapist keeping the session focused on clinically relevant aims, ii) the therapist developing a collaborative relationship with the client, and iii) the therapist making efforts to convey warmth and understanding of the client's thoughts and feelings. In MET session one a rating of 'extensively' was often given to these items, whilst in MET session three, a rating of 'a good deal' was frequently assigned to these items. The frequency of these ratings was unsurprising considering that these items were selected for the BATS for two main reasons. Firstly, to capture the 'common factors' relating to the therapeutic alliance and collaboratively agreeing therapeutic goals. As discussed in the introduction, the therapeutic relationship is considered to be an important common factor (Wampold, 2001), where therapists that possess the ability to form good therapeutic alliances are deemed to achieve more positive outcomes (Wampold, 2015). Secondly, these items were also selected to reflect the key features of therapies widely used in the treatment of alcohol and drug use. For example, both Cognitive Behavioural Therapy (CBT) and Motivational approaches aim to develop a collaborative therapeutic relationship with clients to establish goals in therapy (Beck et al., 1979; Rollnick & Miller, 1995).

There are five basic principles that underlie the MET model (Miller & Rollnick, 1991). These are the need to: express empathy, develop discrepancy, avoid argumentation, roll with resistance, and support self-efficacy. It makes sense therefore that the therapists who delivered the MET intervention in UKATT would aim to develop warm, collaborative relationships with their clients. Consequently, it is unsurprising that BATS items one to three were consistently rated higher than other items on the BATS and perhaps reflects some of these underlying principles of the MET approach.

Furthermore, it is also possible that these items were frequently rated highly due to my project selecting its sample of therapists from UKATT, which was a Randomised Controlled Trial (RCT). As previously highlighted, RCTs often recruit therapists for psychotherapy trials that are experienced and go on to receive training on the therapy model that is being researched in the trial. In doing so, this is likely to lead to low variability in therapist competence and potentially in outcome (Brown et al., 2013). In UKATT, the therapists who delivered the MET intervention all had at least two years of experience in the addiction field. They also received training in the MET intervention and were assessed for competence prior to being signed off to practice in the trial. This potentially reflects the limited variability in these BATS item scores. Consequently, this raises the question as to whether the sample in my study is likely to have been different from therapists in general routine practice, where it is usual for therapists to have a range of experience, professional backgrounds and training. We could expect therefore, a wider range of scores on these items within general routine clinical practice where therapists may be at different stages of their career and are not required to adhere to the strict protocols that exist in RCTs.

Regarding other aspects of the MET approach, it was surprising that item number nine on the BATS ('Considering inconsistencies') was rarely rated during the MET sessions that I observed. As discussed, one of the key underlying principles of the MET approach includes developing discrepancy between the clients' goals or values and their current behaviour (Miller & Rollnick, 1991). The BATS aims to capture this element through item nine, which refers to the therapist exploring how the client's behaviour conflicts with their personal goals or values. I therefore expected to rate this item more frequently in the sessions I observed.

The version of MET used in UKATT was an adapted version of that used in Project MATCH (Miller et al., 1992). Though I did not have access to the UKATT MET manual, I was able to access the Project MATCH MET manual. In the Project MATCH manual, it was suggested that in order to develop discrepancy the therapist could raise the client's awareness of the personal consequences of their drinking behaviour, with the view that this information can create motivation for change. Under developing discrepancy, the manual did not explicitly state that the therapist should relate this to, for example, values, and perhaps explains why I did not rate this BATS item as much as would have been expected.

The items on the BATS that related to 'homework' both being planned with the clients and reviewed by the therapist (items six and seven), were rarely scored. I expected that towards the end of MET session one, homework would have been planned by the therapist collaboratively with their clients. It would have been plausible to expect to have observed the therapist review any previously planned homework at the beginning of MET therapy session three. The findings however suggested that both of these aspects rarely occurred during either of the two sessions.

One of the general ideas in many psychotherapy approaches is that practice of skills in between appointments (often classed as 'homework') can be a valuable therapeutic tool helping to generalise benefits of therapy sessions. This has been associated with positive clinical outcomes of therapy (e.g. Mausbach et al., 2010). Indeed, homework is considered an essential part of some therapies such as CBT (Beck et al., 1979) and Acceptance and Commitment Therapy (Hayes et al., 2012). One possible explanation for BATS items related to homework rarely being scored in the sessions that I observed is due to the adaptation of MET that the therapists were delivering in the trial. As discussed, although I did not have

access to the UKATT MET manual, I was able to access to the Project MATCH manual, which on further exploration did not specify the need for therapists to assign or review homework with clients during sessions. It is likely therefore that this was not a requirement for the therapists delivering the MET intervention within UKATT, which would explain the low ratings across homework items on the BATS in my study.

It is important to acknowledge that questions may arise as to whether the Project MATCH MET manual was significantly different to the UKATT MET manual. However, according to the UKATT researchers, the Project MATCH MET manual was adapted in two ways. Firstly, the MET intervention was limited to three sessions rather than four as in Project MATCH. Secondly, in UKATT significant others were only invited to attend the first MET session with the client, whereas in Project MATCH significant others were able to attend all sessions. This was potentially due to the MET intervention in UKATT being compared to Social Behaviour and Network Therapy (SBNT) as a second active treatment. SBNT is a therapy in which significant others were explicitly invited to attend all therapy sessions and perhaps explains why the BATS item 12 (referring to 'Support Network') was rarely rated during the observation of the MET tapes.

The above findings highlight that when using a transtheoretical measure such as the BATS, not all items will be relevant for all therapeutic models used in the treatment of alcohol and drug use problems. For example, if a therapist is adhering to a manualised intervention that does not require homework tasks, like in UKATT, then these items may not be rated on the BATS. It is important therefore that this is taken into consideration when scoring and interpreting a therapist's performance using a transtheoretical measure.

Furthermore, BATS item eleven which relates to 'Change Planning' was very rarely scored during MET session one, compared to MET session three. This was unsurprising considering the stage of change that most clients were in at the start of therapy. In most cases clients were ambivalent about change or had only just started thinking about changing their alcohol use. This highlights that items on the BATS are not always relevant for every therapy session, an issue that was acknowledged by the author of the BATS and reflects the challenges of developing a scale, which is relevant for different therapies and different stages of the therapy process. Consequently, focus had been given to the overall presentation and structure of the BATS. Its items were grouped based on different stages of change being; i) items considered relevant to most sessions irrespective of client readiness to change, ii) items applicable for building on motivation to change, and iii) items applicable to planning or maintaining change. During the development of the BATS this was considered a viable option to ensure that the BATS was suitable for clinicians to use in routine practice. Alternative designs that target different stages of therapy had been considered but may have increased the length and complexity of the BATS and reduced its utility. Based on my current study it is important to note that there were some instances whereby the 'Change Planning' item (item eleven) was scored during the first MET session. This potentially reflects the stage of change for those individual clients and indicates the usefulness of having the option to score this item when required, providing support for the developer's solution.

4.3. The relationship between the BATS and therapy outcome

Fidelity measures are deemed to be crucial in enabling researchers to develop accurate and meaningful interpretations of the effects of interventions, particularly for the purpose of research. There has however been an increasing need for the use of fidelity measurement in

routine clinical practice (Schoenwald, & Garland, 2013), whereby the delivery of evidence-based treatments is considered vital (Manual et al., 2011). There is the potential for fidelity measures that are designed for and utilised within routine practice to enhance therapist skill and inform clinical supervision.

In routine clinical practice within alcohol and drug use services, therapists often utilise a range of therapies that are tailored to meet the needs of individual clients (Norcross & Wampold, 2011). In order to reflect this practice, a transtheoretical fidelity measure is needed, and thus the BATS was developed (Crosby, 2018). The BATS was designed to evaluate therapist delivery of evidence-based therapies for alcohol and drug use problems with an aim to enhance therapist skill and ultimately clinical outcomes. To provide validation of the BATS, the primary research question in my thesis project aimed to investigate the relationship between the BATS and therapy outcome through the use of the UKATT data (UKATT Research Team, 2005). More specifically, the project aimed to examine whether therapist BATS scores could predict the outcome of therapy.

In my project it was hypothesised that there would be a relationship between the BATS scores and the outcome of the therapy (MET) in UKATT. It was thought that if therapists demonstrate the skills that are frequently considered necessary and important to the delivery of therapies for alcohol and drug use problems, as measured by the BATS, then this would relate to therapeutic change. The analysis revealed that the therapist BATS scores for both MET sessions one and three did not predict the patient's outcome of therapy for either the client's Drinks per Drinking Day (DDD) data or their abstinence outcome data from UKATT.

The absence of a relationship between the BATS scores and outcome of therapy in UKATT in my study is in line with previous research. As detailed in the introduction of my thesis, previous research has sought to examine the relationship between fidelity measures and therapy outcome, often in the context of research trials. What is apparent from the literature is that there are conflicting findings about this relationship. Some studies have identified that decreased fidelity can lead to decreased therapeutic change (e.g. Holder et al., 2018), whilst others have either found a weaker relationship (e.g. Barber et al., 2007) or none at all (e.g. Farmer et al., 2017). However, some research that has specifically examined the link between fidelity and substance use outcomes has found more supportive results regarding this relationship (e.g. Martino et al., 2008). These inconsistent findings reflect the complexity of investigating what makes therapy effective. Frequently in the literature, when researchers study fidelity, this is often in terms of assessing fidelity to a specific model, whilst the BATS assesses fidelity to what are considered effective therapeutic behaviours.

There are potential reasons as to why my study did not identify any relationship between the BATS and client outcomes in UKATT. As previously discussed, exploring the relationship between fidelity measures and therapy outcome is complex. There are many variables other than therapist factors that can influence therapy outcome. Multiple process variables may contribute to the effect of fidelity on outcome by masking or mediating significant associations (Webb et al., 2010). These variables are not always measured in studies and may explain the inconsistencies in research (Perepletchikova & Kazdin, 2005). Indeed, in my study there were many other variables that were not measured or captured in the study or by the BATS measure. These include, the severity of the client's alcohol use, client motivation, readiness to change and the therapeutic alliance (as reported by the therapist or client). These are all factors that can affect therapy outcome (Hogue & Dauber, 2013), and it remains

unclear as to what impact these variables had on patient outcome in my study. It is possible that these factors may have influenced therapy outcome in some way. For example, I recall that upon viewing the MET session recordings, a number of clients showed ambivalence about changing their alcohol consumption at the start of their MET sessions. One of the main aims of the MET approach is to help individuals to resolve their ambivalence. Whilst the therapists may have used effective therapeutic behaviours (as captured by the BATS) that enabled clients to progress towards considering changing their alcohol use, not all clients may have actually changed their drinking behaviours in this brief, three session intervention.

Another potential explanation for the findings in my project may relate to my statistical analysis. I reviewed two therapy sessions for thirty-six patients in UKATT, equating to 72 therapy recordings delivered by a total of thirteen therapists. However, in my analysis I included all patients that each of the thirteen therapists delivered MET therapy to in UKATT – totalling 321, by assigning each therapist's mean BATS scores to each patient that they saw. On reflection, it is unlikely that a therapist would perform consistently across patients, particularly in routine practice, as there is likely to be session variability. There may be factors which affect a therapist's delivery of an intervention, for instance, there may be days whereby a therapist is unable to display as much empathy as others, or some therapists may perform better early in therapy compared to at the end of therapy. There is also the impact of the interaction between therapists and different clients. For example, therapists may adapt their behaviour to the context of treatment, including patient behaviours and characteristics (Stiles, 2009). Consequently, there may be sessions whereby therapists would score higher or lower on the BATS when compared to other sessions. Although I did calculate a mean BATS score for each therapist, the number of sessions observed per therapist was relatively small to the number of cases they had. Indeed, there were three therapists that I only observed once,

and I was therefore unable to calculate a mean BATS score for their sessions. It may have been helpful to have analysed the correlation between the BATS scores for each individual session and the outcome for each of those corresponding patients. It is possible that this may tell us more about the relationship between the BATS and therapy outcome.

Furthermore, as discussed previously the sample of therapists in my study were participating in an RCT. Therapists practicing in an RCT are often selected, trained and monitored to high levels of competence and adherence. This may have therefore restricted the range of scores on the BATS, creating a ceiling effect, which potentially makes it more difficult to determine a true relationship between fidelity and outcome. If I had recruited a sample of therapists based in routine clinical practice in addiction services, and not participating in an RCT, it is possible that there may be more variability in therapists (for example from different training backgrounds and experience) and therefore in BATS scores. This raises the question as to whether we would find a relationship between BATS scores and therapy outcome in research that was conducted within a routine practice setting.

4.4. Therapist variance

In my thesis, the secondary research question aimed to investigate whether the therapists' BATS scores could explain any of the variance in therapist outcomes. One of the key findings of my study was that the BATS scores might explain a portion of the therapist variance in client outcomes. This finding highlighted that the BATS measure may be sensitive enough to capture elements of therapy that are related to the variance in client outcomes and therefore provide a useful tool for therapists in routine clinical practice.

Multilevel modelling revealed that around 2.1% of the variance in client outcomes was attributable to the therapists when I did not adjust for any of the therapist BATS scores or patient level error variance. This percentage of therapist variance changed depending on: i) the outcome that was being used in the analysis (abstinence or DDD), ii) whether the therapist variance was explained by the BATS scores and patient level error variance. When adjusting for the BATS scores, my analysis suggested that the BATS scores might explain a relatively large portion of this 2.1% of unadjusted therapist variance. This key finding is discussed in more detail later in this section.

The finding that around 2.1% of variance in outcomes was attributable to the therapists prior to adjusting for BATS scores was interesting. Relatively few previous studies have produced estimates of this statistic, and those that did have identified modest to large proportions of variability in outcome attributable to the therapist (e.g. Crits-Christoph & Mintz, 1991; Huppert et al., 2001; Project MATCH Research Group, 1997; Wampold & Brown 2005). For example, in a reanalysis of data from ten clinical trials, Crits-Christoph & Mintz (1991) found that the proportion of variance due to the therapists ranged from 0% to 13.5%. Furthermore, in a reanalysis of the National Institute of Mental Health Treatment of Depression Collaborative Research Program (NIHM TDCRP) data through the use of multilevel modelling, Kim et al. (2006) found that around 8% of the variance in outcomes in psychotherapy conditions was attributable to the therapist. However, in their study they found that this percentage of variance ranged depending on the outcome measure that was being used, in addition to the way in which the therapist variance was modelled. This was consistent with the findings in my project whereby the percentage of variance changed depending on the outcome measure that was used in the analysis e.g. DDD or abstinence data.

What is clearly highlighted in the literature is that there are a range of findings relating to therapist effects, with those from my study being relatively small when compared to previous research (e.g. Crits-Christoph & Mintz, 1991; Kim et al., 2006). This raises the question as to why different studies have identified differing levels of therapist variance, and why my project identified a level of variance much smaller in comparison to others.

Wampold & Brown (2005) state that one potential reason for the range of therapist variance reported in studies may be due to whether therapists are considered to be fixed or random factors in analyses. They argue that one of the key difficulties in comparing the results from studies that have investigated therapist effects is that researchers often use various statistical methods, whereby therapists are treated as random or fixed factors in the model. When therapists are considered as fixed effects in the analysis, the results are conditioned on the therapists that were included in the clinical trial, for example in UKATT (Siemer & Joormann, 2003). This increases the level of power to test for main effects. However, by treating therapists instead as random factors i.e. randomly selected from a population of therapists, it is argued that more informative results can be found (Wampold & Brown, 2005). When therapists are considered as being randomly selected from a population of therapists, conclusions can be made about the therapists in general. Indeed, in the aforementioned studies that have investigated therapist effects, the researchers treated therapists as differing factors. For instance, Huppert et al. (2001) considered therapists to be fixed factors in their analyses and identified therapist effect sizes ranging from 0% to 18%. However, Kim et al. (2006) propose that when therapists are treated as random effects, as they are in my study, around 8% of the variance in outcomes is attributable to the therapists.

In my study the percentage of therapist variance, prior to adjusting for the therapist BATS scores, changed depending on the outcome measure that was used in the analysis. Similar findings have been reported in previous research (e.g. Kim et al., 2006). When examining the therapist variance for outcomes using the DDD data, the variance attributable to the therapist was approximately 2.1%. When examining the therapist variance for outcomes using the abstinence data, the therapist variance was around 0%. One potential reason for this may relate to the difference between these outcome measures and the influence of client treatment goals.

Controlled drinking as a client goal rather than abstinence has long been a debated issue in alcohol research. Its use as a treatment goal in clinical practice varies considerably depending on the setting and the country of the addiction service (Luquiens et al., 2011). In the years preceding the UKATT trial, surveys of UK alcohol treatment services identified widespread acceptance of moderate drinking in the UK (Rosenberg et al., 1992). In the years during UKATT, it was found by Cox et al. (2004), that controlled drinking as opposed to complete abstinence was more readily accepted in the UK healthcare system, than compared to other countries in their study. Considering this context at the time of UKATT, it makes sense therefore that the UKATT researchers included both controlled drinking and abstinence as treatment goals in the research design. It is possible that more patients in UKATT opted for controlled drinking and went on to achieve this in comparison to abstinence from alcohol. It would have been interesting to examine this further, however there was no data available in the UKATT database for me to investigate client goals during the trial. Furthermore, although I had collected information on client goals as part of my observations of MET therapy sessions, I was unable to gain access to this data at the time of completing my analysis. This was due to completed BATS score sheets being securely held at the University of Leeds and

inaccessible because of the national lockdown in relation to COVID-19. However, on reflection, my sense was that a greater portion of clients whose sessions I observed reported that their goal for therapy was to control their drinking consumption rather than achieve abstinence. This may explain why there was a larger therapist effect for the DDD outcome measure. Regardless of this however, it is interesting that although the therapist variance was around 0% for the abstinence outcome data, the pattern of results when I adjusted for the therapist BATS scores was similar to that of the DDD data.

One of the key findings regarding the therapist variance identified in my study was that when I adjusted for the therapist BATS scores in the model, the unadjusted therapist variance reduced. This suggested that the BATS scores might explain a portion of the therapist variance. This indicated that the BATS captures elements of therapy that we expect are important therapeutic activities, that are related to the variance in client outcomes.

When adjusting for the BATS scores from the first MET therapy session, the total therapist variance reduced slightly more than when accounting for BATS scores from the third MET session. This potentially suggests that therapist behaviours in the first session were slightly more important than those that were observed in session three. This is consistent with the widely acknowledged importance of the first session of therapy for building rapport and the therapeutic alliance. The therapeutic alliance, deemed to be an important common factor, continues to be one of the most investigated factors that can lead to success in psychotherapy. In a recent meta-analysis of 295 independent studies, researchers confirmed the robustness of a positive relationship between alliance and therapy outcomes (Flückiger et al., 2018). In the literature on the treatment for alcohol problems there have been similar findings. For example, as part of the data that was gathered during Project MATCH, Connors et al. (1997)

examined the relationship between the therapeutic alliance, treatment engagement, and drinking outcomes. Ratings of the working alliance collected from an outpatient sample both during and after treatment (from both the client and the therapist), were found to be significant predictors of client participation in treatment and drinking behaviour, both during treatment and 12-months post-treatment. The first session of therapy is likely to be one whereby the therapist and client are building this working alliance. Indeed, in UKATT, the first MET session was structured around the therapist providing feedback, eliciting the clients concerns and building rapport. This may therefore explain why a greater portion of therapist variance found in this project was explained by the first BATS score derived from MET session one.

The analysis suggested that the BATS was able to explain a relatively large portion of the therapist variance in client outcomes found in this study. However, as discussed previously, the therapist variance found prior to adjusting for the BATS scores was smaller than the therapist variance reported in other studies, at 2.1%. It is possible that the level of therapist variance that was identified in this study is an underestimation, and therefore the proportion of therapist variance that can be explained by the BATS scores is also an underestimation. This is because my sample size was both small and narrow. There were only thirteen therapists that I observed and rated, all of which were recruited into an RCT, received training in MET, and were observed by researchers frequently throughout the trial. This raises the question as to whether we may find more of an effect in a routine clinical practice setting, whereby therapists are not practicing in a tightly controlled trial, and have a range of experience, training and professional backgrounds. Indeed, it has been argued that therapist effects found in naturalistic settings are likely to be greater than in those in RCTs (Wampold, 2015).

Wampold & Brown (2005) aimed to investigate the proportion of variability in outcomes that were attributable to therapists in a managed care setting. This was in order to compare this variance with that which was often found in clinical trials. Their results found that around 5% of variance in outcomes was due to the therapists, when taking into consideration the initial severity of the client. This was a lower value than the therapist variance reported in other studies, which investigated the therapist variance in clinical trials, as reported earlier in this discussion. Wampold & Brown (2005) argued that this was possibly due to there being less patient error variance in clinical trials as clients were more homogeneous than those seen in routine practice. However, in UKATT there was little screening out of clients, which possibly means that the sample of clients seen in UKATT and therefore used in this study is relatively representative of those in routine practice. I would hypothesise, therefore, that there may be more of an effect for both the therapist variance and variance explained by the BATS scores found in routine clinical practice. However, further testing in this area would be required in order to support this hypothesis.

4.5. Reliability of the BATS scoring

The Inter-Rater Reliability (IRR) analysis revealed strong agreement between me (KT) and my supervisor (HC) when scoring the therapists' performance during the first and final MET sessions using the BATS. There were rare instances of disagreement, however any disagreement between our ratings were adjacent e.g. a score of 4 ('extensively') and a score of 3 ('a good deal'). On reflection, the consistency between our ratings was possibly due to me having received training sessions in scoring the BATS with the author of the measure (HC). Clearly defining the individual items on the BATS and discussing how to score them

was useful. This suggests that following training in the use of the BATS, therapist behaviours are likely to be rated similarly if carried out by a different investigator or clinician. This also implies that similar results may be obtained on replication.

4.6. Clinical implications and future research directions

This thesis project was a validation study of the BATS, a transtheoretical measure specifically designed to evaluate the delivery of psychological therapies for substance use problems in routine practice. The BATS scores did not predict client outcome, i.e. higher ratings on the BATS were not associated with a better outcome of therapy. However, analysis did reveal that the BATS might explain some of the therapist variance in outcomes. This indicated that the BATS was able to capture some of the elements of therapy that are deemed to be important therapeutic activities. Consequently, we can assume that the BATS is potentially a useful tool for therapists in both training and supervision in routine clinical practice. Indeed, the BATS has already been incorporated in an NHS addiction service, and permission was granted for addiction services in both Wales and Estonia to use the measure.

The BATS is unusual in that it is not limited to one therapeutic approach or model. The BATS is transtheoretical and can be used across a range of therapeutic approaches often utilised in alcohol and drug use therapy services. This function of the BATS, in addition to the promising findings for this measure outlined earlier in my thesis, warrants the need for further investigation into its potential as a measure to enhance therapist skill in clinical practice. Future studies may benefit from exploring the implementation of the BATS within clinical practice, with a focus on whether therapists' BATS scores correlate with the

outcomes seen in routine practice. Furthermore, as previously mentioned it would be interesting to examine whether the BATS could explain a greater portion of the therapist variance in a clinical setting where the sample of therapists' competence and skill is likely to be broader than that in an RCT.

The BATS includes items that are deemed to be important therapeutic behaviours. As discussed in the introduction of my thesis, previous research into the 'common factors' has examined their relationship with therapy outcome. For example, the development of a collaborative therapeutic alliance has been related to client outcome (e.g. Wampold, 2015; Meier et al., 2005). In addition to this, research has highlighted the potential for this common factor to have complex moderating effects on the relationship between adherence and outcome (Barber et al., 2006). It may be interesting for future research to investigate the relationship between the individual items on the BATS and their relationship with outcome, as this may tell us more about the impact of each of the items.

In my study the BATS was treated as unidimensional whereby the total BATS score for each therapist was included in the analysis. However, it is possible that the BATS is in fact multifactorial, in that it has more than one dimension. If the BATS is indeed made up of more than one dimension, it is also possible that one of those dimensions may be more related to therapeutic outcome than another. This highlights the need for future research to explore the factor structure of the BATS as this will enable us to distinguish whether this transtheoretical fidelity measure is unidimensional or multifactorial. It may therefore be helpful for further research to conduct a factor analysis to investigate the unidimensionality of the BATS. This may provide an understanding of how the twelve items on the BATS relate to one another, which in turn may help us to explain some of the findings in my study.

Whilst my study provides initial support for the reliability of the BATS, it may be useful for future research to build on this by examining the reliability of the BATS with a larger sample, as this is likely to increase the precision of the results (Bowling, 2014). It is also acknowledged that a limited number of sessions were double rated to produce IRR analyses in my study. It would be helpful for future studies to complete a larger portion of sessions which are double rated. Testing other forms of reliability and validity that were not examined in my study would be useful for the ongoing development of the BATS, for example test retest reliability to examine consistency of the BATS scores over time.

4.7. Strengths and Limitations

My thesis project is valuable because it has provided some support for the validation of the BATS whilst investigating an area of research that is not well researched in clinical psychology. By using secondary data from a large RCT, I had access to high quality data. This included both therapy session recordings acquired as part of UKATT and robust outcome data. Though my sample size was relatively small, I was able to observe a significant portion of therapy sessions within the Doctorate in Clinical Psychology time frame.

A key strength of this project relates to the involvement of the author of the BATS (Crosby, 2018). Their input in the project was vital in enabling me to gain a thorough understanding of the development of the BATS, in addition to receiving in depth training in how to score this transtheoretical fidelity measure. Furthermore, it was helpful to have the author of the BATS providing the inter-rater scores for the IRR analyses. Another key strength of my study was that I had access to a statistician who had previous experience in investigating therapist

effects in RCTs. Consequently, this enabled me to explore potential complex analyses, resulting in multilevel modelling which may not have been feasible had I not consulted with a statistician.

Whilst this project has a number of strengths it is not without its limitations. Secondary analyses of data have their advantages in that they are often low-cost projects, and can reduce recruitment difficulties, but they also have disadvantages. It is inherent in the nature of secondary analyses of existing data, that the available data was not collected to address the specific research question of the current study. Consequently, there is less control over the psychotherapy model, the participants, and the quality of session tapes. Furthermore, a major limitation of secondary analyses is that I was not involved in the original data collection process. Therefore, I was unaware of study-specific aspects or anomalies in the data collection process. For example, due to the lack of access to 'completed case' tapes, my sample was not randomised, as initially intended. This resulted in me observing MET sessions from the 'completed cases' that I was able to find in the dataset. Despite which, a strength in the chosen design for my study was that it made this piece of research achievable in the Doctorate in Clinical Psychology time frame.

One limitation referred to throughout the discussion is that the number of therapists used in this project was small and may represent a skewed population, which may not represent therapists in routine clinical settings. The therapists in UKATT were therapists working in specialist addiction services, but all had to demonstrate a level of competence to participate in the trial. The therapists were trained in the MET protocol and were monitored throughout their involvement in the RCT. It is therefore plausible that the MET intervention delivered in

the UKATT study was of higher quality than can be expected from uncontrolled community settings.

Finally, it is important to consider the utility of the BATS response scale, which may be questioned as a potential limitation of the measure. The BATS enables the rater to score each item on a five point Likert scale that ranges from 'not at all' to 'extensively'. A strength of Likert scales is that they allow a range of responses to be quantified. Although the BATS provides five choices of response, it could be argued that these scales are subjective and present choices that can be difficult to determine between, for example being able to clearly differentiate between conveying empathy (item three) 'extensively' or 'a good deal'. Furthermore, it has been argued that the space between each choice on a Likert scale cannot be equidistant, meaning that the scale may fail to capture the true attitude of the rater (Bishop & Herron, 2015). These points indicate that there is likely to be some differences in opinion regarding ratings assigned by different raters. Consideration was given to how the responses would be scored by the author of the BATS (Crosby, 2018). The author consulted with experts in the field on how items on the BATS should be rated, specifically on whether these should be rated on an extensiveness or quality scale. The participant feedback in the original BATS development study (Crosby, 2018) suggested that using a Likert scale to rate the extensiveness of therapist behaviours was an appropriate option, as this would make the items easier and more reliable to rate in comparison to a quality scale. Rating a scale based on quality would require judgements of competence and would be harder to sufficiently measure. Further support for the chosen response scale for the BATS includes the IRR ratings in my study, which demonstrated good to excellent levels of agreement between raters.

4.8. Conclusions

This thesis has sought to examine the validity of the BATS, which is a transtheoretical fidelity measure for monitoring and evaluating therapists' delivery of psychological therapies in routine practice for alcohol and drug use problems. In being a transtheoretical measure, the BATS is applicable to the range of widely used therapies in addiction services. The BATS was unable to predict therapy outcome in my sample. This finding was consistent with previous literature in this area. It is a reminder of the complexity of therapy, and the many variables that influence therapy outcome. Although the BATS was unable to predict therapy outcome, an interesting finding was that my analysis of therapist variance in client outcomes suggested that the BATS was able to explain a proportion of the therapist variance. This indicates that the BATS was sensitive enough to capture some of the elements widely considered to be important in therapy. Furthermore, the BATS has demonstrated excellent levels of inter-rater reliability. Overall, my project suggests that the BATS is a useful tool for training and supervision of therapists in substance use services, where it may help to enhance therapists' competence and treatment outcomes.

References

Ball, S. A., Martino, S., Corvino, J., Morganstern, J., & Carroll, K. M. (2002). *Independent tape rater guide.* Unpublished psychotherapy tape rating manual.

Barber, J. P., & Crits-Christoph, P. (1996). Development of an adherence/competence scale for dynamic therapy: Preliminary findings. *Psychotherapy Research, 6,* 81−94. http://dx.doi.org/10.1080/10503309612331331608

Barber, J. P., Gallop, R., Crits-Christoph, P., Frank, A., Thase, M. E., Weiss, R. D., & Connolly Gibbons, M. B. (2006). The role of therapist adherence, therapist competence, and alliance in predicting outcome of individual drug counselling: Results from the National Institute Drug Abuse Collaborative Cocaine Treatment Study. *Psychotherapy Research, 16*(2), 229-240. https://doi.org/10.1080/10503300500288951

Barber, J. P., Liese, B. S., & Abrams, M. J. (2003). Development of the cognitive therapy adherence and competence scale. *Psychotherapy Research, 13*(2), 205-221. https://doi.org/10.1093/ptr/kpg019

Barber, J. P., Mercer, D., Krakauer, I., & Calvo, N. (1996). Development of an Adherence/Competence Rating Scale for individual drug counseling. *Drug and Alcohol Dependence, 43,* 125-132. https://doi.org/10.1016/s0376-8716(96)01305-1

Barber, J. P., Sharpless, B. A., Klostermann, S., & McCarthy, K. (2007). Assessing intervention competence and its relation to therapy outcome: A selected review derived from the outcome literature. *Professional Psychology: Research and Practice, 38*(5). https://doi.org/10.1037/0735-7035.38.5.493

Bassett, S. S., Stein, L. A. R., Rossi, J. S., & Martin, R. A. (2016). Evaluating measures of fidelity for substance abuse group treatment with incarcerated adolescents. *Journal of Substance Abuse Treatment. 66,* 9-15. https://doi.org/10.1016/j.jsat.2016.02.011

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

Beck, A. T., Ward, C. H., Medelson, M., Mock, J., & Erbaugh, J. (1961). An inventory for measuring depression. *Archives of General Psychiatry, 4*, 561-571. https://doi.org/10.1001/archpsyc.1961.01710120031004

Bishop, P. A., & Herron, R. L. (2015). Use and misuse of the Likert item responses and other ordinal measures. *International Journal of Exercise Science, 8*(3), 297-302.

Bjaastad, J. F., Haugland, B. S. M., Fjermestad, K. W., Torsheim, T., Havik, O. E., Heiervang, E. R., & Ost, L. (2016). Competence and Adherence Scale for Cognitive Behavioural Therapy (CAS-CBT) for anxiety disorders in youth: Psychometric properties. *Psychological Assessment, 28*(8), 908-916.

Bjaastad, J. F., Wergeland, G. J. H., Haugland, B. S. M., Gjestad, R., Havik, O. E., Heiervang, E. R., & O, L. (2018). Do clinical experience, formal cognitive behavioural therapy training, adherence, and competence predict outcome in cognitive behavioural therapy for anxiety disorders in youth? *Clinical Psychology and Psychotherapy*, *25*, 865-877. https://doi.org/10.1002/cpp.2321

Black, S., Hardy, G., Turpin, G., & Parry, G. (2005). Self-reported attachment styles and therapeutic orientation of therapists and their relationship with reported general alliance quality and problems in therapy. *Psychology and Psychotherapy*, *78*(3), 363-377. https://doi.org/10.1348/147608305x43784

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

Borrelli, B. (2011). The assessment, monitoring and enhancement of treatment fidelity in public health clinical trials. *Journal of Public Health Dentistry, 71*, S52-S63. https://doi.org/10.1111/j.1752-7325.2011.00233.x

Borrelli, B., Sepinwall, D., Ernst, D., Bellg, A., Czajkowski, S., Breger, R., DeFrancesco, C., Levesque, C., Sharp, D. L., Ogedegbe, G., Resnick, B., & Orwig, D. (2005). A new tool to access treatment fidelity and evaluation treatment fidelity across 10 years of health behaviour research. *Journal of Consulting and Clinical Psychology, 73*(5), 852-860. https://doi.org/10.1037/0022-006X.73.5.852

Branson, A., Myles, P., Mahdi, M., & Shafran, R. (2018). The relationship between competence and patient outcome with low-intensity cognitive behavioural interventions. *Behavioural and Cognitive Psychotherapy, 46*, 101-114. https://doi.org/10.1017/S1352465817000522

Branson, A., Shafran, R., & Myles, P. (2015). Investigating the relationship between competence and patient outcome with CBT. *Behaviour Research and Therapy, 68*, 19-26. https://doi.org/10.1016/j.brat.2015.03.002

Breitenstein, S. M., Gross, D., Garvey, C. A., Hill, C., Fogg, L., & Resnick, B. (2010). Implementation fidelity in community-based interventions. *Research in Nursing & Health. 33*(2), 164-173. https://doi.org/10.1002/nur.20373

Brown, L. A., Craske, M. G., Glenn, D. E., Stein, M. B., Sullivan, G., Sherbourne, C., Bystritsky, A., Welch, S. S., Campbell-Sills, L., Lang, A., Roy-Byrne, P., & Rose, R. D. (2013). CBT competence in novice therapists improves anxiety outcomes. *Depression and Anxiety, 30*, 97-115. https://doi.org/10.1002/da.22027

Brown, S. D., & Lent, R. W. (2008). *Handbook of Counselling Psychology* (4th ed.). John Wiley & Sons.

Campos-Melady, M., Smith, J. E., Meyers, R. J., Godley, S. H., & Godley, M. D. (2017). The effect of therapists' adherence and competence in delivering the adolescent community reinforcement approach on client outcomes. *Psychology of Addictive Behaviours, 31*(1), 117-129. https://doi.org/10.1037/adb0000216

Carroll, K. M. (1997). *Cognitive-Behavioral Coping Skills Treatment for Cocaine Dependence*. https://medicine.yale.edu/psychiatry/pdc/Images/CBT_tcm678-175259.pdf

Carroll, K. M. (1998). *A Cognitive Behavioral Approach: Treating Cocaine Addiction.* National Institute on Drug Abuse. https://archives.drugabuse.gov/sites/default/files/cbt.pdf

Carroll, K. M., Connors, G. J., Cooney, N. L., DiClemente, C. C., Donovan, D. M., Kadden, R. R., Longabaugh, R. L., Rounsaville, B. J., Wirtz, P. W. & Zweben, A. (1998). Internal validity of Project MATCH treatments: Discriminability and integrity. *Journal of Consulting & Clinical Psychology. 66*(2), 290-303. https://doi.org/10.1037/0022-006X.66.2.290

Carroll, K. M., Martino, S. & Rounsaville, B. J. (2010). No train, no gain? *Clinical Psychology: Science and Practice. 17*(1), 36-40. https://doi.org/10.1111/j.1468-2850.2009.01190.x

Carroll, K. M., Nich, C., Ball, S. A., McCance, E., Frankforter, T. L., & Rounsaville, B. J. (2000). One-year follow-up of disulfiram and psychotherapy for cocaine-alcohol users: Sustained effects of treatment. *Addiction, 95*(9), 1335-1349. https://doi.org/10.1046/j.1360-0443.2000.95913355.x

Chow, D. L., Miller, S. D., Seidel, J. A., Kane, R. T., Thornton, J. A., & Andrews, W. P. (2015). The role of deliberate practice in the development of highly effect psychotherapists. *Psychotherapy, 52*(3), 337-345. http://dx.doi.org/10.1037/pst0000015

Cicchetti, D. V. (1994). Guidelines, criteria, and rules of thumb for evaluating normed and standardized assessment instruments in psychology. *Psychological Assessment, 6,* 284–290. https://doi.org/10.1037/1040-3590.6.4.284

Collyer, H., Eisler, I., & Woolgar, M. (2019). Systematic literature review and meta-analysis of the relationship between adherence, competence and outcome in psychotherapy for children and adolescents. *European Child and Adolescent Psychiatry*, *29*(5). https://doi.org/10.1007/s00787-018-1265-2

Copello, A., Orford, J., Hodgson, R., Tober, G., & Barrett, C. (2002). Social behaviour and network therapy: Basic principles and early experiences. *Addictive Behaviors, 27*(3), 345-366. https://doi.org/10.1016/s0306-4603(01)00176-9

Connors, G. J., Carroll, K. M., DiClemente, C. C., Longabaugh, R., & Donovan, D. M. (1997). The therapeutic alliance and its relationship to alcoholism treatment participation and outcome. *Journal of Consulting and Clinical Psychology, 65*(4), 588-598. https://doi.org/10.1037//0022-006x.65.4.588

Cox, W. M., Rosenberg, H., Hodgins, H. A., Macartney, J. I., & Maurer, K. A. (2004). United Kingdom and United States healthcare providers' recommendations of abstinence versus controlled drinking. *Alcohol & Alcoholism, 39*(2), 130-134. https://doi.org/10.1093/alcalc/agh035

Crits-Christoph, P., & Mintz, J. (1991). Implications of therapist effects for the design and analysis of comparative studies of psychotherapies. *Journal of Consulting and Clinical Psychology, 59,* 20-26. https://doi.org/10.1037//0022-006x59.1.20

Crits-Christoph, P., Siqueland, L., Blaine, J., Frank, A., Luborsky, L., Onken, L.S., Muenz, L. R., Thase, M. E., Weiss, R. D., Gastfriend, D. R., Woody, G. E., Barber, J. P., Butler, S. F., Daley, D., Salloum, I., Bishop, S., Njavits, L. M., Lis, J., Mercer, D., . . . Beck. A. T. (1999). Psychosocial treatments for cocaine dependence: National Institute on Drug Abuse Collaborative Cocaine Treatment Study. *Archives of General Psychiatry, 56*(6), 493-502. https://doi.org/10.1001/archpsyc.56.6.493

Crosby, H. F. (2018). *Development of the Brief Addiction Therapist Tool (BATS): A tool for evaluating therapist delivery of psychological therapies for alcohol and drug use problems* [Unpublished doctoral thesis]. University of Leeds.

Culbreth, J. R. (1999). Clinical supervision of substance abuse counsellors: Current and preferred practices. *Journal of Addictions and Offender Counselling, 20,* 15-25. https://doi.org/10.1002/j.2161-1874.1999.tb00137.x

Currat, T., Despland, J. N., Drapeau, M., Beretta, V., Fischer, M., & de Roten, Y. (2008). How many years are required for developing competence for a very brief form of psychotherapy? A pilot study. *Swiss Archives of Neurology, Psychiatry and Psychotherapy, 159,* 358-364. https://doi.org/10.4414/sanp.2008.01977

Dane, A. V., & Schneider, B. H. (1998). Program integrity in primary and early secondary prevention: Are implementation effects out of control? *Clinical Psychology Review, 18*(1), 23-45. https://doi.org/10.1016/S0272-7358(97)00043-3

David, D., Cristea, I., & Hofmann, S. G. (2018). Why cognitive behavioural therapy is the current gold standard of psychotherapy. *Front Psychiatry, 9*(4). https://doi.org/10.339/fpsyt.2018.00004

Degenhardt, L., Charlson, F., Mathers, B., Hall, W. D., Flaxman, A. D., Johns, N., & Vos, T. (2014). The global epidemiology and burden of opioid dependence: results from the global burden of disease 2010 study. *Addiction, 109*(8), 1320-1333. https://doi.org/10.1111/add.12551

Despland, J., Roten, Y., Drapeau, M., Currat, T., Beretta, V., & Kramer, U. (2009). The role of alliance in the relationship between therapist competence and outcome in brief psychodynamic psychotherapy. *The Journal of Nervous and Mental Disease, 197*(5), 362-367. https://doi.org/10.1097/NMD.0b013e3181a20849

DiClemente, C. C., & Prochaska, J. O. (1998). Toward a comprehensive, transtheoretical model of change: Stages of change and addictive behaviors. In W. R. Miller & N. Heather (Eds.), *Applied clinical psychology. Treating addictive behaviors* (pp. 3–24). Plenum Press.

Elkin, I. (1999). A major dilemma in psychotherapy outcome research: Disentangling therapists from therapies. *Clinical Psychology Science and Practice, 6*(1), 10-32. https://doi.org/10.1093/clipsy.6.1.10

Emmelkamp, P. M. G., David, D., Beckers, T., Muris, P., Cuijpers, P., Lutz, W., Andersson, G., Araya, R., Banos Rivera, R. M., Barkham, M., Berking, M., Berger, T., Botella, C., Carlbring, P., Colom, f., Essau, C., Hermans, D., Hofmann, S. G., Knappe, S., . . . Vervliet, B. (2014). Advancing psychotherapy and evidence-based psychological interventions. *International Journal of Methods in Psychiatric Research, 23*(1), 58-91. https://doi.org/10.1002/mpr.1411

Falender, C. A., & Shafranske, E. P. (2004). *Clinical supervision: A competency-based approach.* American Psychological Association.

Farmer, C. C., Mitchell, K. S., Parker-Guilbert, K., & Galovski, T. E. (2017). Fidelity to the cognitive processing therapy protocol: Evaluation of critical elements. *Behaviour Therapy, 48*, 195-206. https://doi.org/10.1016/j.beth.2016.02.009

Flückiger, C., Del Re, A. C., Wampold, B. E., & Horvath, A. O. (2018). The alliance in adult psychotherapy: A meta-analytic synthesis. *Psychotherapy, 55*(4), 316-340. http://dx.doi.org/10.1037/pst0000172

Folke, S., Daniel, S. I. F., Gondan, M., Lunn, S., Taekker, L., & Poulsen, S. (2017). Therapist adherence is associated with outcome in cognitive behavioural therapy for bulimia nervosa. *Psychotherapy, 54*(2), 195-200. https://doi.org/10.1037/pst0000107

Frank, E., Kupfer, D. J., Wagner, E. F., McEachran, A. B. & Cornes, C. (1991). Efficacy of interpersonal therapy as a maintenance treatment of recurrent depression. *Archives of General Psychiatry 48*, 1053–1059. https://doi.org/10.1001/archpsyc.1991.01810360017002

Fureman, I., Parikh, G., Bragg, A., & McLellan, A. T. (1990). *Addictions Severity Index* (5th ed.). University of Pennsylvania Veterans Administration Centre for Addiction.

Gearing, R. E., El-Bassel, N., Ghesquiere, A., Baldwin, S., Gillies, J., & Ngeow, E. (2011). Major ingredients of fidelity: A review and scientific guide to improving quality of intervention research implementation. *Clinical Psychology Review, 31*(1), 79-88. https://doi.org/10.1016/j.cpr.2010.09.007

Gloster, A. T., Einsle, F., Lang, T., Hauke, C., & Wittchen, H. (2008). *TACRS: The Therapist Adherence and Competency Rating Scales or panic disorder and agoraphobia.* Unpublished instrument.

Hagermoser Sanetti, L. M., Collier-Meek, M. A., & Fallon, L. M. (2016). Fidelity with Flexibility: Treatment Acceptability and Individualised Adaptations of Evidence-Supported Treatments. In S. Maltzman, *The Oxford handbook of treatment process and outcomes in psychology: A multidisciplinary, biopsychosocial approach* (pp. 292). Oxford University Press.

Haug, T., Nordgreen, T., Ost, L., Tangen, T., Kvale, G., Hovland, O. J., Heiervang, E. R., & Havik, O. E. (2016). Working alliance and competence as predictors of outcome in cognitive behavioural therapy for social anxiety and panic disorder in adults. *Behaviour Research and Therapy, 77*, 40-51. https://doi.org/10.1016/j.brat.2015.12.004

Hauke, C., Gloster, A. T., Gerlach, A. L., Richter, J., Kircher, T., Fehm, L., Stoy, M., Lang, T., Klotsche, J., Einsle, F., Deckert, J., & Wittchen, H. U. (2014). Standardized treatment

manuals: Does adherence matter? *Sensoria: A Journal of Mind, Brain & Culture, 10*(2), 1–13. https://doi.org/10.7790/sa.v0i0.362

Hayes, S. C. (1998). Market-driven treatment development. *The Behavior Therapist, 21*, 32–33.

Hayes, S. C., Strosahl, K. D., & Wilson, K. G. (2012). *Acceptance and commitment therapy: The process and practice of mindful change* (2nd ed.). The Guilford Press.

Hogue, A., & Dauber, S. (2013). Assessing fidelity to evidence-based practices in usual care: The example of family therapy for adolescent behavior problems. *Evaluation and Program Planning, 37,* 21–30. https://doi.org/10.1016/j.evalprogplan.2012.12.001

Hogue, A., Liddle, H. A., & Rowe, C. (1996). Treatment adherence process research in family therapy: A rationale and some practical guidelines. *Psychotherapy: Theory, Research, Practice, Training, 33*(2), 332-345. https://doi.org/10.1037/0033-3204.33.2.332

Holder, N., Holliday, R., Williams, R., Mullen, K., & Surís, A. (2018) A preliminary examination of the role of psychotherapist fidelity on outcomes of cognitive processing therapy during an RCT for military sexual trauma-related PTSD. *Cognitive Behaviour Therapy, 47*(1), 76-89. https://doi.org/10.1080/16506073.2017.1357750

Hollon, S. D., Evans, M. D., Auerbach, A., DeRubeis, R. J., Elkin, I., & Lowery, A. (1988). *Development of a system for rating therapies for depression: Differentiating cognitive*

therapy, interpersonal psychotherapy and clinical management pharmacotherapy. Unpublished manuscript, Vanderbilt University, Nashville.

Horvath, A. O., & Greenberg, L. S. (1989). Development and validation of the Working Alliance Inventory. *Journal of Counseling Psychology, 36*, 223–233.

Huppert, J. D., Barlow, D. H., Gorman, J. M., Shear, M. K., & Woods, S. W. (2006). The Interaction of Motivation and Therapist Adherence Predicts Outcome in Cognitive Behavioral Therapy for Panic Disorder: Preliminary Findings. *Cognitive and Behavioral Practice, 13*(3), 198–204. https://doi.org/10.1016/j.cbpra.2005.10.001

Huppert, J. D., Bufka, L. F., Barlow, D. H., Gorman, J. M., Shear, M. K., & Woods, S. W. (2001). Therapists, therapist variables, and cognitive behavioural therapy outcome in a multicentre trial for panic disorder. *Journal of Consulting and Clinical Psychology, 69*(5), 747-755. https://doi.org/10.1037//0022-006x.69.5.747

IBM Corp. (2015). *IBM SPSS Statistics for Windows* (Version 23.0). IBM Corp.

Kedzior, K. K., & Tabata Laeber, L. (2014). A positive association between anxiety disorders and cannabis use or cannabis use disorders in the general population – a meta-analysis of 31 studies. *BMC Psychiatry, 14*(136). https://doi.org/10.1186/1471-244X-14-136

Kim, D., Wampold, B. E., & Bolt, M. (2006). Therapist effects in psychotherapy: A random-effects modelling of the National Institute of Mental Health Treatment of Depression

Collaborative Research Program data. *Psychotherapy Research, 16*(2), 161-172. http://dx.doi.org/10.1080/10503300500264911

Lambert, M. J., & Barley, D. E. (2002). Research summary on the therapeutic relationship and psychotherapy outcome. In J. C. Norcross (Eds.), *Psychotherapy Relationships that Work: Therapist Contributions and Responsiveness to Patients* (pp. 17-32). Oxford University Press.

Liddle, H. (2010). Treating adolescence substance abuse using multidimensional family therapy. Evidence-based psychotherapies for children and adolescents. In J. Weisz., & A. Kazdin (Eds.), *Evidence Based Psychotherapies for Children and Adolescents* (pp. 416-432). The Guilford Press.

Liu, J., Tang, W., Chen, G., Lu, Y., Feng, C., & Tu, X. M. (2016). Correlation and agreement: overview and clarification of competing concepts and measures. *Shanghai Archives of Psychiatry, 28*(2), 115-120.

Loeb, K. L., Wilson, G. T., Labouvie, E., Pratt, E. M., Hayaki, J., Walsh, B. T., & Fairburn, C. G. (2005). Therapeutic alliance and treatment adherence in two interventions for bulimia nervosa: A study of process and outcome. *Journal of Consulting and Clinical Psychology, 73*, 1097– 1107. http://dx.doi.org/10.1037/0022-006X.73.6.1097

Loucks, S. F. (1983). *Defining fidelity: A cross-study analysis* [Paper presentation]. Annual Meeting of the American Educational Research Association, Montreal, Quebec, Canada.

Luquiens, A., Reynaud, M., & Aubin, H. J. (2011). Is controlled drinking an acceptable goal in the treatment of alcohol dependence? A survey of French alcohol specialists. *Alcohol and Alcoholism, 46*(5), 586-591. https://doi.org/10.1093/alcalc/agr083

Madson, M. B., & Campbell, T. C. (2006). Measures of fidelity in motivational enhancement: A systematic review. *Journal of Substance Abuse Treatment, 31*(1), 67-73. https://doi.org/10.1016/j.jsat.2006.03.010

Manderscheid, R. W. (1998). From many to one: Addressing the crisis of quality in managed behavioral health care at the millennium. *The Journal of Behavioral Health Services & Research, 25*, 233–238. https://doi.org/10.1007/bf02287484

Manuel, J. K., Hagedorn, H. J., & Finney, J. W. (2011). Implementing evidence-based psychosocial treatment in speciality substance use disorder care. *Psychology of Addictive Behaviours, 25*(2), 225-237. https://doi.org/10.1037/a0022398

Martino, S., Ball, S. A., Nich, C., Frankforter, T. L., & Carroll, K. M. (2008). Community program therapist adherence and competence in motivational enhancement therapy. *Drug and Alcohol Dependence, 96*, 37-48. https://doi.org/10.1016/j.drugalcdep.2008.01.020

Mausbach, B. T., Moore, R., Roesch, S., Cardenas, V., & Patterson, T. L. (2010). The relationship between homework compliance and therapy outcomes: An updated meta-analysis. *Cognitive Therapy and Research, 34*(5), 429-438. https://doi.org/10.1007/s10608-010-9297-z

McCambridge, J., Day, M., Thomas, B., & Strang, J. (2012). Fidelity to motivational interviewing and subsequent cannabis cessation among adolescents three months after brief intervention. *Addiction Science & Clinical Practice, 7*(Suppl 1), A8. https://doi.org/10.1186/1940-0640-7-S1-A8

McDonnell, M. G., Rodgers, M. L., Short, R. A., Norell, D., Pinter, L., & Dyck, D. G. (2007) Clinician integrity in multiple family groups: psychometric properties and relationship with schizophrenia client and caregiver outcomes. *Cognitive Therapy and Research, 31*, 785-803

Meier, P. S., Barrowclough, C., & Donmall, M. C. (2005). The role of the therapeutic alliance in the treatment of substance misuse: a critical review of the literature. *Addiction, 100*(3), 304-316. https://doi.org/10.1111/j.1360-0443.2004.00935.x

Merlo, J., Chaix, B., Ohlsson, H., Beckman, A., Johnell, K., Hierpe, P., Rastam, L., & Larsen, K. (2006). A brief conceptual tutorial of multilevel analysis in social epidemiology: using measures of clustering in multilevel logistic regression to investigate contextual phenomena. *Journal of Epidemiology and Community Health, 60*(40), 290-297. https://doi.org/10.1136/jech.2004.029454

Middleton, W., Tober, G., Frier, E., & Finnegan, O., with the UKATT Process Data Sub-Group. (2001). *UKATT Process Rating Manual*. Unpublished manual.

Miller, W. R. (1996). *Form 90: A structured assessment interview for drinking and related behaviors.* (Project MATCH monograph series, vol. 5). National Institute on Alcohol Abuse and Alcoholism.

Miller, S. J. & Binder, J. L. (2002). The effects of manual based training on treatment fidelity and outcome: a review of the literature on individual adult psychotherapy. *Psychotherapy: Theory, Research, Practice, Training, 39*, 184-198. https://doi.org/10.1037/0033-3204.39.2.184

Miller, W. R., & Rollnick, S. (1991). *Motivational Interviewing: Preparing people to change addictive behaviour.* The Guilford Press.

Miller, W. R., & Rollnick, S. (2013). *Motivational Interviewing: Helping People Change* (3rd ed.). The Guilford Press.

Miller, W. R., Zweben, A., DiClemente, C. C. & Rychtarik, R. (1992). *Motivational Enhancement Therapy manual: A clinical research guide for therapists treating individuals with alcohol abuse and dependence*. (Project MATCH Monograph Series, Vol. 2). National Institute on Alcohol Abuse and Alcoholism.

Miller, W. R., Zweben, J., & Johnson, W. R. (2005). Evidence-based treatment: why, what, where, when, and how? *Journal of Substance Abuse Treatment, 29*(4), 267-276. https://doi.org/
10.1016/j.jsat.2005.08.003

Moncher, F., & Prinz, R. J. (1991). Treatment fidelity in outcome studies. *Clinical Psychology Review, 11*(13), 247-266. https://doi.org/10.1016/0272-7358(91)90103-2

Moyers, T., Martin, T., Manuel, J. K., Hendrickson, S. M., & Miller, W. R. (2005). Assessing competence in the use of motivational interviewing. *Journal of Substance Abuse Treatment, 28*(1), 19−26. https://doi.org/10.1016/j.jsat.2004.11.001

NHS Digital. (2019). *Statistics on Alcohol, England 2019: Alcohol-related Hospital Admissions.* NHS Digital. https://digital.nhs.uk/data-and-information/publications/statistical/statistics-on-alcohol/2019/part-1

Norcross, J. C., & Wampold, B. E. (2011). Evidence-based therapy relationships: Research conclusions and clinical practices. *Psychotherapy, 48*(1), 98–102. https://doi.org/10.1037/a0022161

O'Malley, S. S., Foley, S. H., Rounsaville, B. J., Watkins, J. T., Sotsky, S. M., Imber, S. D., & Elkin, I. (1988). Therapist competence and patient outcome in interpersonal psychotherapy of depression. *Journal of Consulting and Clinical Psychology, 56*(4), 496–501. https://doi.org/10.1037/0022-006X.56.4.496

Perepletchikova, F., & Kazdin, A. E. (2005). Treatment integrity and therapeutic change: Issues and research recommendations. *Clinical Psychology Science and Practice, 12*(4), 365-383. https://doi.org/10.1093/clipsy.bpi045

Perepletchikova, F., Treat, T. A., & Kazdin, A. E. (2007). Treatment integrity in psychotherapy research: Analysis of the studies and examination of the associated factors.

Journal of Consulting and Clinical Psychology, 75, 829-841. https://doi.org/10.1037/0022-006X.75.6.829

Petry, N. M. & Stitzer, M. L. (2002). *Contingency Management: Using Motivational Incentives to Improve Drug Abuse Treatment.* Yale University Psychotherapy Development Center.

Project MATCH Research Group. (1997). Matching Alcoholism Treatments to Client Heterogeneity: Project MATCH posttreatment drinking outcomes. *Journal of Studies on Alcohol. 58*(1), 7-29. https://doi.org/10.15288/jsa.1997.58.7

Prowse, P., & Nagel, T. (2015). A meta-evaluation: The role of treatment fidelity within psychosocial interventions during the last decade. *Journal of Psychiatry, 18*(2). https://doi.org/10.4172/2378-5756.1000251

Raistrick, D., Heather, N., & Godfrey, C. (2006). *Review of the effectiveness of treatment for alcohol problems.* National Treatment Agency for Substance Misuse. http://www.dldocs.stir.ac.uk/documents/alcoeffective.pdf

Rapley, H. A., & Loades, M. E. (2019). A systematic review exploring therapist competence, adherence and therapy outcomes in individual CBT for children and young people. *Psychotherapy Research, 29*(8), 1010-1019. https://doi.org/10.1080/10503307.2018.1464681

Richards, D., & Whyte, M. (2009). *Reach Out: National programme student materials to support the delivery of training for Psychological Wellbeing Practitioners delivering low intensity interventions* (2nd ed.). Rethink.

Rollnick, S., & Miller, W. R. (1995). What is motivational interviewing? *Behavioural and Cognitive Psychotherapy*, *23*, 325-334. https://doi.org/10.1017/S135246580001643X

Rosenberg, H., Melville, J., Levell, D., & Hodge, J. (1992). A 10-year follow-up survey of acceptability of controlled drinking in Britain. *Journal of Studies on Alcohol 53*, 441–446. https://doi.org/10.15288/jsa.1992.53.441

Rosenzweig, S. (1936). Some implicit common factors in diverse methods of psychotherapy. *American Journal of Orthopsychiatry, 6*(3), 412-415. https://doi.org/10.1111/j.1939-0025.1936.tb05248.x

Santacroce, S. J., Maccarelli, L. M., & Grey, M. (2004). Intervention fidelity. *Nursing Research, 53*(1), 63-66.

Schoenwald, S. K., & Garland, A. F. (2013). A review of treatment adherence measurement methods. *Psychology Assessment, 25*(1), 146-156. https://doi.org/10.1037/a0029715

Schoenwald, S. K., Garland, A. F., Chapman, J. E., Frazier, S. L., Sheidow, A. J., & Southam-Gerow, M. A. (2011). Toward the effective and efficient measurement of implementation fidelity. *Administration and Policy in Mental Health, 38*(1), 32-43. https://doi.org/10.1007/s10488-010-0321-0

Shapiro, D. A., & Shapiro D. (1982). Meta-analysis of comparative therapy outcome studies: a replication and refinement. *Psychology Bulletin, 92*(2), 581-604. https://doi.org/10.1037//0003-066X.32.9.752

Shaw, B. F., Elkin, I., Yamaguchi, J., Olmsted, M., Vallis, T. M., Dobson, K. S., Lowery, A., Sotsky, S. M., Watkins, J. T., & Imber, S. D. (1999). Therapist competence ratings in relation to clinical outcome in cognitive therapy of depression. *Journal of Consulting and Clinical Psychology, 67*(6), 837-846. https://doi.org/10.1037//0022-006x.67.6.837

Shrout, P. E., & Fleiss, J. L. (1979). Intraclass correlations: Uses in assessing rater reliability. *Psychological Bulletin, 86*, 420–429. https://doi.org/10.1037//0033-2909.86.2.420

Siemer, M., & Joorman, J. (2003). Power and measures of effect sizes in analysis of variance with fixed versus random nested factors. *Psychological Methods, 8*, 497-517. https://doi.org/10.1037/1082-989X.8.4.497

Smith, M. L., & Glass, G. V. (1977). Meta-analysis of psychotherapy outcome studies. *American Psychologist, 32*(9), 752–760. https://doi.org/10.1037//0003-066X.32.9.752

Snijders, T.A.B. & Bosker, R.J. (1999). *Multilevel analysis: An Introduction to basic and advanced multilevel modelling* (Eds.). Sage Publications Inc.

Sprenkle, D. H., & Blow, A. J. (2004). Common factors and our sacred models. *Journal of Martial and Family Therapy, 30*(2), 113-129. https://doi.org/10.1111/j.1752-0606.2004.tb01228.x

Sprenkle, D. H., Davis, S. D., & Lebow, J. L. (2013). *Common factors in couple and family therapy: The overlooked foundation for effective practice paperback.* The Guilford Press.

StataCorp. (2017). *Stata Statistical Software: Release 15*. StataCorp LLC.

Stiles, W. B. (2009). Responsiveness as an obstacle for psychotherapy outcome research: It's worse than you think. *Clinical Psychology: Science and Practice, 16*(1),86-91.

Streiner, D. L., Norman, G. R., & Cairney, J. (2015). *Health measurement scales: A practical guide to their development and use* (5th ed.). Oxford University Press.

Strunk, D. R., Brotman, M. A., & DeRubeis, R. J. (2010). The process of change in cognitive therapy for depression: predictors of early inter-session symptom gains. *Behaviour Research and Therapy, 48*(7), 599-606. https://doi.org/10.1016/j.brat.2010.03.011

Tadic, M., Drapeau, M., Solai, S., de Roten, Y., & Despland, J. N. (2003). Development of a competence scale for brief psychodynamic investigation: A pilot study. *Swiss Archives for Neurology, Psychiatry and Psychotherapy, 154*, 28-36.

Tober, G., Clyne, W., Finnegan, O., Farrin, A., Russell, I., & UKATT Research Team. (2008). Validation of a scale for rating the delivery of psycho-social treatments for alcohol dependence and misuse: The UKATT Process Rating Scale (PRS). *Alcohol and Alcoholism. 43*(6), 675-682. https://doi.org/10.1093/alcalc/agn064

Tober, G., & Crosby, H. F. (2011). *AESOPS Process Rating Manual Version 5*. Unpublished manual.

Tober, G., & Crosby, H. F. (2014). *ADAPTA Process Rating Scale Manual*. Unpublished manual.

Tober, G., Kenyon, R., Brodie, J., Heather, N., & Raistrick, D. (2002). *Manual for Motivational Enhancement Therapy*. Unpublished manual.

Tonigan, J. S., Miller, W. R., & Brown, J. M. (1997). The reliability of Form 90: An instrument for assessing alcohol treatment outcome. *Journal of Studies on Alcohol, 58*(4), 358–364. https://doi.org/10.15288/jsa.1997.58.358

Torrey, A. R. (2012). Treatment fidelity instrument to measure a brief opportunistic intervention for prenatal substance use. *Journal of Obstetric, Gynaecologic, and Neonatal Nursing, 41*(6), 809-820. https://doi.org/10.1111/j.1552-6909.2012.01413.x

Trepka, C., Rees, A., Shapiro, D. A., Hardy, G. E., & Barkham, M. (2004). Therapist competence and outcome of cognitive therapy for depression. *Cognitive Therapy and Research, 28*(2), 143-157. https://doi.org/10.1023/B:COTR.0000021536.39173.66

Trevethan, R. (2017). Intraclass correlation coefficients: clearing the air, extending some cautions, and making some requests. *Health Services and Outcomes Research Methodology, 17*(2), 127-143. https://doi.org/10.1007/s10742-016-0156-6

Truax, C. B., & Carkhuff, R. R. (1967). *Toward effective counseling and psychotherapy: Training and practice*. Aldine Publishing Co.

UKATT Research Team. (2001). United Kingdom alcohol treatment trial (UKATT): Hypotheses, design and methods. *Alcohol & Alcoholism 36*(1), 11-21. https://doi.org/10.1093/alcalc/36.1.11

UKATT Research Team. (2005). UKATT: Effectiveness of treatment for alcohol problems: findings of the randomised United Kingdom Alcohol Treatment Trial (UKATT). *British Medical Journal, 331*, 541-544. https://doi.org/10.1136/bmj.331.7516.541

Vasilaki, E. I., Hosier, S. G., & Cox, W. M. (2006). The efficacy of motivational interviewing as a brief intervention for excessive drinking: a meta-analytic review. *Alcohol Alcohol, 41*(3), 328-335. https://doi.org/10.1093/alcalc/agl016

Waltz, J., Addis, M. E., Koerner, K., & Jacobson, N. S. (1993). Testing the integrity of a psychotherapy protocol: Assessment of adherence and competence. *Journal of Consulting and Clinical Psychology, 61*(4), 620-630. https://doi.org/10.1037//0022-006x.61.4.620

Wampold, B. E. (2001). *The great psychotherapy debate: Models, methods, and findings.* Lawrence Erlbaum Associates Publishers.

Wampold, B. E. (2015). How important are the common factors in psychotherapy? An update. *World Psychiatry, 14*(3), 270-277. https://doi.org/10.1002/wps.20238

Wampold, B. E., & 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*(5), 914-923. https://doi.org/10.1037/0022-006X.73.5.914

Wampold, B. E., & Imel, Z. E. (2015). *The great psychotherapy debate: The evidence for what makes psychotherapy work* (2nd ed.). Routledge/Taylor & Francis Group.

Wampold, B. E., & Serlin, R. C. (2000). The consequence of ignoring a nested factor on measures of effect size in analysis of variance. *Psychological Methods, 5*(4), 425 – 433. https://doi.org/10.1037/1082-989X.5.4.425

Watson, J., Crosby, H., Dale, V., Tober, G., Wu, Q., Lang, J., McGovern, R., Newbury-Birch, D., Parrott, S., & Bland, J. (2013). AESOPS: a randomised controlled trial of the clinical effectiveness and cost-effectiveness of opportunistic screening and stepped care interventions for older hazardous alcohol users in primary care. *Health Technology Assessment, 17*(25), 1-158. https://doi.org/10.3310/hta17250

Watson, J.M., Fairhurst, C., Li, J., Tober, G., Crosby, H., Lloyd, C., Godfrey, C., Mdege, N.D., Dale, V., Toner, P., Parrott, S., & Raistrick, D. (2015). ADAPTA: A pilot randomised controlled trial of an alcohol-focused intervention versus a healthy living intervention for problem drinkers identified in a general hospital setting. *Drug and Alcohol Dependence, 154*, 117-124. https://doi.org/10.1016/j.drugalcdep.2015.06.030

Webb, C.A., DeRubeis, R.J., & Barber, J.P. (2010). Therapist adherence/competence and treatment outcome: A meta-analytic review. *Journal or Consulting and Clinical Psychology, 78*(2), 200-211. https://doi.org/10.1037/a0018912

Weck, F., Hautzinger, M., Heidenreich, T., & Stangier, U. (2011). Psychoeducation for depression. Features of interventions and therapeutic competencies. *Psychotherapie, Psychosomatik, Medizinische Psychologie, 61*(3-4), 148-53. https://doi.org/10.1055/s-0030-1269902

Weck, F., Weigel, M., Hautzinger, M., Barocka, A., Schlosser, R. G. M., & Stangier, U. (2012). Relapses in recurrent depression 1 year after psychoeducational treatment: The role of therapist adherence and competence, and the therapeutic alliance. *Psychiatry Research, 195*(1-2), 51-55. https://doi.org/10.1016/j.psychres.2011.07.025

Weigal, M., Weck, F., Rudari, V., Hilling, C., Hautzinger, M., Heidenreich, T., & Stangier, U. (2009). *The Manualised Active Psychoeducation – Adherence Scale (MAPE-AS).* Unpublished manuscript, University of Frankfurt, Germany.

Weisman, A., Tompson, M. C., Okazaki, S., Gregory, J., Goldstein, M. J., Rea, M., & Miklowitz, D.J. (2002). Clinician's fidelity to a manual-based family treatment as a predictor of the one-year course of bipolar disorder. *Family Processes, 41*, 123–131. https://doi.org/10.1111/j.1545-5300.2002.40102000123.x

Willig, C. (2013). *Introducing Qualitative Research in Psychology* (3rd ed.). McGraw-Hill Education

World Health Organisation. (2018). *Alcohol*. https://www.who.int/news-room/fact-sheets/detail/alcohol

Yeaton, W., & Sechrest, L. (1981). Critical dimensions in the choice and maintenance of successful treatments: Strength, integrity, and effectiveness. *Journal of Consulting and Clinical Psychology, 49*, 156–167. https://doi.org/10.1037//0022-006X.49.2.156

Appendix A: The Brief Addiction Therapist Scale (BATS)

B A T S

Brief Addiction Therapist Scale

A tool for evaluating therapists' delivery of psychological therapies for alcohol and drug use problems.

Designed to facilitate training and supervision, and enhance therapist skill.

Instructions for use: For each item, circle a number on the 5-point scale reflecting the extent to which the therapist carried out the behaviour. For items that are not applicable to the session, score 0 'not at all'. Use the space provided on page 2 to give context, comments, and additional information e.g. the client's first session. Item definitions are provided on page 2. To be used with audio or video recordings of therapy sessions.

During the session...	Not at all	A little	Somewhat	A good deal	Extensively
1. The therapist kept the session focused on the aims for that session.	0	1	2	3	4
2. The therapist attempted to work together with the client.	0	1	2	3	4
3. The therapist conveyed empathy.	0	1	2	3	4
4. The therapist focused on the client's strengths.	0	1	2	3	4
5. The therapist used "complex reflections" – offering a perspective which added meaning and enabled the client to make connections.	0	1	2	3	4
6. The therapist and the client planned tasks for the client to do between sessions.	0	1	2	3	4
7. The therapist and the client reviewed tasks planned in the previous session.	0	1	2	3	4
8. The therapist enabled the client's goals for treatment to be discussed.	0	1	2	3	4
If in this session the focus was on building motivation for change:					
9. The therapist encouraged the client to consider inconsistencies between their substance use, and personal goals or values.	0	1	2	3	4
10. The therapist encouraged the client to talk about the positive aspects of changing substance use.	0	1	2	3	4
If in this session the focus was on planning or maintaining change:					
11. The therapist enabled a plan for changing the client's substance use, or maintaining change, to be discussed.	0	1	2	3	4
12. The therapist discussed how the client's social network might support changing substance use or maintaining change.	0	1	2	3	4

Total score:	

Version 1 | bats@leeds.ac.uk

Item definitions

1. **Session aims:** The therapist kept the session focused on clinically relevant aims during the session, e.g. target behaviour. This may or may not include explicit discussion of the purpose of the session, *e.g. to describe a relapse prevention plan*. Aims may change during the course of the session following disclosure of risk.
2. **Working together**: Developing a collaborative relationship between the client and the therapist. It is about discussing, actively seeking the client's input; not telling, and not arguing.
3. **Convey empathy**: Making efforts to convey warmth and understanding of the client's thoughts and feelings. The therapist avoids any blaming or labelling.
4. **The client's strengths**: Helping the client to identify and focus on what they can do, not what they cannot do: achievements rather than failings.
5. **Complex reflections**: Helping the client to gain insight by making and/or strengthening connections between things they have said. Going beyond repeating or slightly rephrasing what the client has said.
6. **Planning tasks**: Any task that is planned (the therapist and the client agreed what to do and how to do it) for the client to do between sessions, *e.g. specific homework tasks, trying new behaviours*.
7. **Reviewing tasks**: Explicit discussion in which tasks set in the previous session are reviewed. This item is not applicable if it is the client's first session, tick the box as appropriate.
8. **Treatment goals**: Goals refer to the overall treatment goals, *e.g. abstinence, harm reduction, moderation*. The goals could be discussed by the therapist and/or the client.
9. **Considering inconsistencies**: Exploring how the client's behaviour conflicts with his/her personal goals and values, *e.g. I need to drink a bottle of gin but I want to be a good parent*.
10. **Talking about change**: The therapist encourages the client to talk about the positive aspects of changing.
11. **Change planning**: Discussion of an overall plan to achieve the agreed treatment goals. Tasks represent the steps in the plan to achieve the overall treatment goals.
12. **The social network**: The therapist facilitates a discussion about the client's actual and/or potential social network to identify how this may support the overall plan.

Context, comments, and additional information:

We acknowledge the support of Leeds and York Partnership NHS Foundation Trust (LYPFT). BATS was funded by Alcohol Research UK (ARUK), and the Society for the Study of Addiction (SSA). We thank the participants involved in developing BATS.

UNIVERSITY OF LEEDS

Appendix B: Literature search strategy

PsycINFO (1806 to 2019)

1. Fidelity ADJ1 Therapist ADJ3 Outcome
2. Fidelity ADJ1 Therapist ADJ3 Change
3. Fidelity ADJ1 Therapist ADJ3 effectiveness
4. Fidelity ADJ1 Therapist ADJ3 therap*
5. Fidelity ADJ1 practitioner ADJ3 Outcome
6. Fidelity ADJ1 practitioner ADJ3 change
7. Fidelity ADJ1 practitioner ADJ3 effectiveness
8. Fidelity ADJ1 practitioner ADJ3 efficacy
9. Fidelity ADJ1 practitioner ADJ3 therap*
10. Fidelity ADJ1 clinician ADJ3 Outcome
11. Fidelity ADJ1 clinician ADJ3 change
12. Fidelity ADJ1 clinician ADJ3 effectiveness
13. Fidelity ADJ1 clinician ADJ3 efficacy
14. Fidelity ADJ1 clinician ADJ3 therap*
15. Therapist ADJ1 adherence ADJ3 outcome
16. Therapist ADJ1 adherence ADJ3 change
17. Therapist ADJ1 adherence ADJ3 effectiveness
18. Therapist ADJ1 adherence ADJ3 efficacy
19. Therapist ADJ1 adherence ADJ3 therap*
20. Practitioner ADJ1 adherence ADJ3 outcome
21. Practitioner ADJ1 adherence ADJ3 change
22. Practitioner ADJ1 adherence ADJ3 effectiveness
23. Practitioner ADJ1 adherence ADJ3 efficacy
24. Practitioner ADJ1 adherence ADJ3 therap*
25. Clinician ADJ1 adherence ADJ3 outcome
26. Clinician ADJ1 adherence ADJ3 change
27. Clinician ADJ1 adherence ADJ3 effectiveness
28. Clinician ADJ1 adherence ADJ3 efficacy
29. Clinician ADJ1 adherence ADJ3 therap*
30. Therapist ADJ1 competence ADJ3 outcome
31. Therapist ADJ1 competence ADJ3 change
32. Therapist ADJ1 competence ADJ3 effectiveness
33. Therapist ADJ1 competence ADJ3 efficacy
34. Therapist ADJ1 competence ADJ3 therap*
35. Practitioner ADJ1 competence ADJ3 outcome
36. Practitioner ADJ1 competence ADJ3 change
37. Practitioner ADJ1 competence ADJ3 effectiveness
38. Practitioner ADJ1 competence ADJ3 efficacy
39. Practitioner ADJ1 competence ADJ3 therap*

40. Clinician ADJ1 competence ADJ 3 outcome
41. Clinician ADJ1 competence ADJ3 change
42. Clinician ADJ1 competence ADJ3 effectiveness
43. Clinician ADJ1 competence ADJ3 efficacy
44. Clinician ADJ1 competence ADJ3 therap*
45. Integrity ADJ1 therapist ADJ3 outcome
46. Integrity ADJ1 therapist ADJ3 change
47. Integrity ADJ1 therapist ADJ3 effectiveness
48. Integrity ADJ1 therapist ADJ3 efficacy
49. Integrity ADJ1 therapist ADJ3 therap*
50. Integrity ADJ1 clinician ADJ3 outcome
51. Integrity ADJ1 clinician ADJ3 change
52. Integrity ADJ1 clinician ADJ3 effectiveness
53. Integrity ADJ1 clinician ADJ3 efficacy
54. Integrity ADJ1 clinician ADJ3 therap*
55. Integrity ADJ1 practitioner ADJ3 outcome
56. Integrity ADJ1 practitioner ADJ3 change
57. Integrity ADJ1 practitioner ADJ3 effectiveness
58. Integrity ADJ1 practitioner ADJ3 efficacy
59. Integrity ADJ1 practitioner ADJ3 therap*

Medline (1946 to 2019)

1. Fidelity ADJ1 Therapist ADJ3 Outcome
2. Fidelity ADJ1 Therapist ADJ3 Change
3. Fidelity ADJ1 Therapist ADJ3 effectiveness
4. Fidelity ADJ1 Therapist ADJ3 therap*
5. Fidelity ADJ1 practitioner ADJ3 Outcome
6. Fidelity ADJ1 practitioner ADJ3 change
7. Fidelity ADJ1 practitioner ADJ3 effectiveness
8. Fidelity ADJ1 practitioner ADJ3 efficacy
9. Fidelity ADJ1 practitioner ADJ3 therap*
10. Fidelity ADJ1 clinician ADJ3 Outcome
11. Fidelity ADJ1 clinician ADJ3 change
12. Fidelity ADJ1 clinician ADJ3 effectiveness
13. Fidelity ADJ1 clinician ADJ3 efficacy
14. Fidelity ADJ1 clinician ADJ3 therap*
15. Therapist ADJ1 adherence ADJ3 outcome
16. Therapist ADJ1 adherence ADJ3 change
17. Therapist ADJ1 adherence ADJ3 effectiveness
18. Therapist ADJ1 adherence ADJ3 efficacy

19. Therapist ADJ1 adherence ADJ3 therap*
20. Practitioner ADJ1 adherence ADJ3 outcome
21. Practitioner ADJ1 adherence ADJ3 change
22. Practitioner ADJ1 adherence ADJ3 effectiveness
23. Practitioner ADJ1 adherence ADJ3 efficacy
24. Practitioner ADJ1 adherence ADJ3 therap*
25. Clinician ADJ1 adherence ADJ3 outcome
26. Clinician ADJ1 adherence ADJ3 change
27. Clinician ADJ1 adherence ADJ3 effectiveness
28. Clinician ADJ1 adherence ADJ3 efficacy
29. Clinician ADJ1 adherence ADJ3 therap*
30. Therapist ADJ1 competence ADJ3 outcome
31. Therapist ADJ1 competence ADJ3 change
32. Therapist ADJ1 competence ADJ3 effectiveness
33. Therapist ADJ1 competence ADJ3 efficacy
34. Therapist ADJ1 competence ADJ3 therap*
35. Practitioner ADJ1 competence ADJ3 outcome
36. Practitioner ADJ1 competence ADJ3 change
37. Practitioner ADJ1 competence ADJ3 effectiveness
38. Practitioner ADJ1 competence ADJ3 efficacy
39. Practitioner ADJ1 competence ADJ3 therap*
40. Clinician ADJ1 competence ADJ 3 outcome
41. Clinician ADJ1 competence ADJ3 change
42. Clinician ADJ1 competence ADJ3 effectiveness
43. Clinician ADJ1 competence ADJ3 efficacy
44. Clinician ADJ1 competence ADJ3 therap*
45. Integrity ADJ1 therapist ADJ3 outcome
46. Integrity ADJ1 therapist ADJ3 change
47. Integrity ADJ1 therapist ADJ3 effectiveness
48. Integrity ADJ1 therapist ADJ3 efficacy
49. Integrity ADJ1 therapist ADJ3 therap*
50. Integrity ADJ1 clinician ADJ3 outcome
51. Integrity ADJ1 clinician ADJ3 change
52. Integrity ADJ1 clinician ADJ3 effectiveness
53. Integrity ADJ1 clinician ADJ3 efficacy
54. Integrity ADJ1 clinician ADJ3 therap*
55. Integrity ADJ1 practitioner ADJ3 outcome
56. Integrity ADJ1 practitioner ADJ3 change
57. Integrity ADJ1 practitioner ADJ3 effectiveness
58. Integrity ADJ1 practitioner ADJ3 efficacy
59. Integrity ADJ1 practitioner ADJ3 therap*

Appendix C: Summary of literature examining fidelity and its relationship with outcome

Table C1: Summary of literature examining the relationship between fidelity and outcome in individual psychotherapy literature since 2008

Author and year	Country	Aims	Clinical Population	Participants	Aspects of Fidelity	Method	Fidelity measure	Analysis	Findings	Limitations
Branson et al. (2018)	UK	To examine the relation between competence and outcome.	Adults presenting with anxiety or depression, treated with CBT.	3688 clients 47 PWPs	Competence	Competence examined through 3 OSCEs, rated using criteria and learning outcomes set in the national curriculum for PWPs (Richards & Whyte, 2009)	None Therapists measured against national curriculum.	Spearman rho correlation between OSCEs 1-3 & reliable change in patient's symptoms.	Little support for general association between competence and patient outcome either during or post-training.	Uncontrolled retrospective nature of the study meant it was not possible to ascertain how accurately data was entered e.g. missing at random or routine failure to input.
Branson et al. (2015)	UK	Examine the relationship between CBT competence and outcome.	Adults with mild to moderate anxiety and/or depression treated with CBT.	1247 clients 43 therapists enrolled on the IAPT CBT training programme.	Competence	Trainee therapist audio recordings of therapy rated by experienced CBT therapists.	Cognitive Therapy Scale Revised (CTS-R) (Blackburn et al., 2001)	Spearman rho correlations between mean CTS-R scores and outcome.	Little support for general association between competence and outcome. Significantly more clients of the most competent therapists	Therapists self-selected sessions for rating potentially resulting in a restriction of range.

									demonstrated reliable improvement. Fewer experienced no reliable change. Significantly more clients of least competent therapists had a reliable deterioration in symptoms.	
Brown et al. (2013)	USA	Examined the relationship between competence and adherence and outcomes of computer-assisted CBT for anxiety disorders.	Adults with moderate to clinically significant anxiety.	176 adults experiencing moderate to clinically significant anxiety. 14 novice therapists.	Adherence Competence	259 therapy sessions (ranging from sessions 1 to 20) recorded, observed and rated.	Measures of CBT adherence and CBT competence were created for the study.	RCIs within multilevel modelling framework. A series of regressions were also conducted.	Higher therapist competence predicted better outcomes at 12 and 18 months. This effect was not present at 6-month follow-up. Therapist adherence did not significantly predict outcome.	Scales used to measure therapist competence and adherence were developed specifically for this study. Scales lacked psychometric development.
Campos-	USA	To examine	Adolescent	384 Clients	Adherence	Independent	The A-	Multilevel	Competence	No measure of

Melady et al. (2017)	ratings of adherence and competence and their relationship with outcome.	substance use.	from the AAFT project 91 Therapists trained to deliver A-CRA	Competence	ratings of audio recorded sessions.	CRA Procedures Checklist.	models.	significantly predictive of decrease in client days of substance use. Adherence not predictive of substance use outcome. When clients completed 12-month follow-up between therapist adherence was predictive of decrease in substance use.	general therapist skill.
Despland et al. (2009)	To examine the role of alliance in the relationship between competence and outcome in BPP.	78 outpatients presenting with a 'psychiatic disorder' (e.g. a mood disorder, anxiety or personality	78 clients treated with 4 sessions of therapy based on psychoanalytic theory. 15 therapists, with training in PP.	Competence	All BPP sessions were videotaped. First sessions analysed for competence by six trained raters.	BPI-CS (Tadic et al., 2003; Currat et al., 2008)	Nested design using hierarchical linear modelling.	No direct link between therapist competence and outcome. Results suggested alliance patterns as a moderator in relationship	Patients were not randomly assigned to therapists. There was no controlled distribution of patients to therapists according to years of

			disorder).						between competence and outcome.	training and experience.
Farmer et al. (2017).	USA	To examine therapist fidelity as a predictor of session to session PTSD and depression symptom change. Therapist competence across course of therapy on symptom change.	Adults with PTSD treated with CPT.	Recruited from a trial examining the efficacy of CPT for interpersonal assault survivors. This study included data from 46 trial completers and 22 dropouts. 8 therapists	Adherence Competence	533 taped CPT sessions.	CPT:TACP-R	Repeated measures regression models using SAS Proc Mixed to examine within and between clients scores as predictors of symptom change. Competence and symptom change analysed using univariate ANOVA.	No significant effects of therapist fidelity on session to session symptom change. High competence related to some elements of CPT components was related to greater client change in PTSD severity. Therapist competence ratings overall were not significantly associated with	Relatively small sample which may limit the ability to generalise findings.

									improvement in depression symptomatology	
Bjaastad et al. (2018).	Norway	To investigate whether clinical experience, formal CBT training, adherence, and competence predict outcome in individual and group CBT for anxiety disorders in youth.	Children and young people with anxiety disorder	182 young people 17 Therapists	Adherence Competence	181 therapy videos observed and rated for adherence and competence.	CAS-CBT for anxiety disorders in youth (Bjaastad et al., 2016).	Latent growth curve modelling.	Higher therapist adherence related to better outcomes. Competence related to worse outcome. Findings were not consistent. Interaction effects suggested that competence amongst therapists with formal CBT training was related to better client outcome.	Level of experience for CBT education and therapist competence in the group condition was only investigated for main therapist and not co-therapist. Potential effects of difference for co-therapist experience were not addressed in the analyses.

Folke et al. (2017)	Denmark	To investigate change in adherence to CBT for bulimia nervosa over time. To investigate the relationship between adherence and outcome in different phases of treatment.	Adults with bulimia nervosa treated with enhanced CBT (20 sessions).	36 clients who received enhanced CBT for bulimia nervosa. 4 therapists with eight years of therapy experience.	Adherence	Trained raters rated 92 audiotapes of full-length therapy sessions from early, middle and late phases of therapy.	CBT:TPAS (Loeb et al., 2005)	Multilevel Poisson regression analysis.	Higher levels of adherence in early and middle phases of treatment were associated with reduced binging frequency. Higher levels of adherence measured late in treatment was not.	The small number of therapists precludes any conclusions to be drawn regarding therapist variability in adherence in the sample.
Haug et al. (2016).	Norway	To examine working alliance, competence and adherence as predictors of outcome for CBT for SAD and PD.	Adults with SAD or PD treated with 12 sessions of face-to-face CBT	82 clients with SAD (*n* = 51) or PD (*n* = 31) 22 Therapists with limited CBT experience.	Adherence Competence	Therapist competence and adherence assessed by independent raters from videos of therapy sessions 3 and 8.	CTACS (Barber et al., 2003).	Unadjusted regression analyses and hierarchical multiple regression analyses.	Higher therapist competence and adherence early in therapy associated with better outcome among PD clients. Lower	The raters were not completely independent as two of the three raters were clinical supervisors and one was a therapist in the project.

									competence and adherence associated with dropout among SAD clients.	
Holder et al. (2018)	USA	An examination of the effect of therapist fidelity during CPT on outcomes during an RCT for military sexual trauma-related PTSD.	Veterans with military related PTSD	72 veterans who received CPT for military related PTSD 4 therapists who delivered CPT.	Adherence Competence	One independent fidelity expert who viewed and rated on average 12% of CPT session videos.	No specific measure to monitor fidelity. Therapists rated against the essential element for CPT sessions.	Hierarchical linear modelling.	Clients treated by therapists with 'good' fidelity had significantly greater reductions in PTSD symptoms, NCs and depression symptoms than clients treated by therapists with 'below' average treatment fidelity.	The inclusion of only four therapists limited the ability to investigate and control for a wide variety of therapist factors.
Hauke et al. (2014).	Germany	To examine therapist adherence across the course of CBT	Adults diagnosed with panic disorder with agoraphobia.	220 adult clients 58 Therapists	Adherence	560 videotapes from baseline to post therapy (session 1 to session 12)	Therapist Adherence and Competency Rating Scales (Gloster et al., 2008)	Mixed-effects Linear regression models.	Adherence did not influence outcome at global outcome. Symptom	Detection of meaningful patterns may have been restricted by the reduced variance in therapist

									severity and patient motivation interact with adherence to predict outcome.	adherence ratings e.g. nearly 80% of sessions were rated as moderately – highly adherent.
Martino et al. (2008)	USA	To examine adherence and competence of interventions associated with MI and general counselling	Adults with substance use problems.	461 outpatients 35 Therapists	Adherence Competence	15 Independent tape raters rated audiotapes of sessions.	ITRS (Ball et al., 2002)	Pearson correlations	Higher levels of MI adherence and competence associated with increases in client motivation and some positive client treatment outcomes in community programs.	Process analysis was only on clients who completed three sessions which limits the variability in outcomes by excluding some clients who dropped out of treatment and may not have had positive outcomes.
McCambridge et al. (2012)	UK	To examine whether differences in cannabis cessation 3 months after MI sessions was	Adolescent (16-19 years) frequent cannabis users.	75 clients 4 Practitioners delivering MI	Adherence Competence	Audio-recording of sessions were rated by two coders.	MITI scale (version 2) Moyers et al. (2005)	Multilevel logistic regression model	Two aspects of enhanced fidelity (MI spirit and complex reflections) were predictive of cannabis	Clients in the study may have accessed additional support during the study which was not accounted for.

		attributable to therapist fidelity.							cessation 3 months after MI session.	
Strunk et al. (2010)	USA	To examine therapist adherence to therapeutic techniques as predictors of symptom improvement in CT for depression.	Adults with moderate to severe depression.	60 clients 6 therapists	Adherence	Tapes rated by two raters trained in CT who were blind to outcome.	Adherence measured by a subset of items from the CSPRS (Hollon et al., 1988).	Regression analyses.	Support for the idea that elements of adherence to methods of CT predict symptom change. Adherence to cognitive methods was the strongest of all predictors examined.	No control or comparison group in the study. Therefore, unsure if findings are specific to CT.
Weck et al. (2012)	Germany	To examine the role of adherence, competence and alliance in relapse of recurrent depression after psychoeducation treatment.	Adults with recurrent depression who received 16 sessions of psychoeducational treatment.	43 clients 40 Therapists trained in the treatment.	Adherence Competence	Video tapes were analysed to evaluate adherence and competence by four judges.	Adherence measured by MAPE-AS (Weigal et al., 2009) Competence measured by CS-P (Weck et al., 2011)	Correlations between adherence, competence, alliance and outcome were analysed using Pearson's correlation coefficient.	One year following treatment, no associations found between adherence and competence and outcome.	Statistical power limited due to only 43 individual treatments being taken into account.

AAFT = Assertive Adolescent and Family Treatment project; **A-CRA** = Adolescent Community Reinforcement Approach; **ANOVA** = Analysis of Variance; **BPI-CS**= Brief Psychodynamic Intervention Competence Scale; **BPP** = Brief Psychodynamic Psychotherapy; **CAS-CBT** = Competence and Adherence Scale for CBT; **CBT** = Cognitive Behavioural Therapy; **CBT:TPAS** = Cognitive-Behavioral Therapy Treatment Protocol Adherence Scale; **CPT** = Cognitive Processing Therapy; **CPT:TACP-R** = Cognitive Processing Therapy: Therapist Adherence and Competence Protocol – Revised; **CS-P** = Competence Scale

for Psychoeducation; **CSPRS** = Collaborative Study Psychotherapy Rating Scale; **CT** = Cognitive Therapy; **CTACS** = Cognitive Therapy Adherence and Competence Scale; **CTS-R** = Cognitive Therapy Scale - Revised; **IAPT** = Improving Access to Psychological Therapies; **ITRS** = Independent Tape Rating Scale; **MAPE-AS** = Manualised Active Psychoeducation – Adherence Scale; **MI** = Motivational Interviewing; **MITI** = Motivational Interviewing Treatment Integrity scale**; NCs** = Negative Cognitions; **OSCE** = Observed Standardised Clinical Examinations; **PD** = Panic disorder; **PP** = Psychodynamic Psychotherapy; **PTSD** = Post-Traumatic Stress Disorder; **PWP** = Psychological Wellbeing Practitioner; **RCI** = Reliable Change Index; **RCT** = Randomised Controlled Trial; **SAD** = Social Anxiety Disorder

Appendix D: Original UKATT consent form

UKATT Participant Consent Form

I agree to take part in the research comparing two forms of help for stopping or reducing drinking.

The research has been explained to me. I understand that I will be offered one of two forms of help and that I will be required to complete some further questionnaires during the therapy and to attend follow-up appointments. I also understand that, with my consent, someone (or more than one) who knows me well may be involved in meetings with the therapist.

I understand that any personal information I give in this research project will be kept strictly confidential. I understand that this information will be used only in combination with information from many other people so that I cannot be identified.

I understand that, with my consent, the member of my family or other person who knows me well whom I have suggested, may be contacted for further information of my progress after the end of the therapy. I understand that any information from this other individual will be kept strictly confidential. I also understand that any other contact names and addresses I have supplied will be used purely for establishing my whereabouts during the follow-up period and my involvement in this trial will not be revealed to them.

I agree to video recordings of my sessions being used for quality control and teaching purposes, and for future research. I understand that I will not be seen in the video but my voice will be heard on the recording. I understand that by putting a cross in the appropriate box below these tapes will be destroyed at the end of the trial.

I know that I can ask questions about the research now or at any stage, and that I can choose to withdraw from the research at any time without this affecting the quality of the help I receive.

I have been given a list of the names and telephone numbers of those responsible for this research, including the name of a manager to whom I should address any complaint or grievance that I might have.

I require that all video recordings of my session be destroyed at the end of the trial ☐

Name……. Assessors Name……..

Signature…… Signature……

Date……

Appendix E: Ethical approval MREC 18-077

The Secretariat
University of Leeds
Leeds, LS2 9JT
Tel: 0113 3431642
Email: FMHUniEthics@leeds.ac.uk

UNIVERSITY OF LEEDS

Kate Thresher
Psychologist in Clinical Training
Clinical Psychology
Leeds Institute of Health Sciences
Faculty of Medicine and Health
Level 10, Worsley Building
University of Leeds
Clarendon Way
LEEDS LS2 9NL

09 May 2019

Dear Kate

Ref no: **MREC 18-077**
Study Title: **Evaluation of the Brief Addiction Therapist Scale (BATS) through the secondary analysis of data**

Thank you for submitting your documentation for the above project. Following review by the School of Medicine Research Ethics Committee (SoMREC) I can confirm a conditional favourable ethical opinion based on the documentation received at date of this letter *and subject to the following condition/s which must be fulfilled prior to the study commencing:*

- **Evidence of gatekeeper approvals to access the UKATT data from the University of York must be submitted prior to the study commencing**
- **Evidence of gatekeeper approvals to access the therapy recordings from the Division of Psychological and Social Medicine, University of Leeds, must be submitted prior to the study commencing**

The study documentation must be amended as required to meet the above conditions and submitted for file and possible future audit. Once you have addressed the conditions and submitted for file/future audit, you may commence the study and further confirmation of approval is not provided.

Please note, failure to comply with the above conditions will be considered a breach of ethics approval and may result in disciplinary action.

Document Received	*Version*	*Date Received*
Ethics Application - Kate Thresher Thesis	1.0	15/03/2019
UKATT original consent form	1.0	15/03/2019

Please notify the committee if you intend to make any amendments to the original research as submitted at date of this approval. This includes recruitment methodology and all changes must be ethically approved prior to implementation. Please contact the Faculty Research Ethics Administrator for further information FMHUniEthics@leeds.ac.uk

Ethical approval does not infer you have the right of access to any member of staff or student or documents and the premises of the University of Leeds. Nor does it imply any right of access to the premises of any other organisation, including clinical areas. The SoMREC takes no responsibility for you gaining access to staff, students and/or premises prior to, during or following your research activities.

You are expected to keep a record of all your approved documentation, as well as documents such as sample consent forms, risk assessments and other documents relating to the study. This should be kept in your study file, and may be subject to an audit inspection. If your project is to be audited, you will be given at least 2 weeks notice.

It is our policy to remind everyone that it is your responsibility to comply with Health and Safety, Data Protection and any other legal and/or professional guidelines there may be.

The committee wishes you every success with your project.

Yours sincerely

Dr Naomi Quinton
Co-Chair, School of Medicine Research Ethics Committee

Appendix F: Ethics amendment approval

RE: MREC 18-077 Am d 1 Sept 2019 - Approval with Caveat

From: Rachel De Souza [Medicine] <R.E.DeSouza@leeds.ac.uk> on behalf of Medicine and Health Univ Ethics Review <FMHUniEthics@leeds.ac.uk>
Sent: 14 October 2019 09:52
To: Kate Thresher <umkt@leeds.ac.uk>
Cc: Medicine and Health Univ Ethics Review <FMHUniEthics@leeds.ac.uk>; h.crosby@leedstrinity.ac.uk <h.crosby@leedstrinity.ac.uk>
Subject: RE: MREC 18-077 Am d 1 Sept 2019 - Approval with Caveat

Hi Kate

MREC 18-077 Amd 1 Sept 2019 - Evaluation of the Brief Addiction Therapist Scale (BATS) through the secondary analysis of data

Many thanks for submitting the above amendment. I can confirm this has been approved on the understanding that an encrypted memory stick is used and a caveat that the file names on the encrypted memory stick must not contain any identifiers (i.e. names or locations).

Please retain this email as evidence of approval in your study file.

Please notify the committee if you intend to make any amendments to the information in your ethics application as submitted at date of this approval as all changes must receive ethical approval prior to implementation. The amendment form is available at http://ris.leeds.ac.uk/EthicsAmendment.

Please note: You are expected to keep a record of all your approved documentation and other documents relating to the study, including any risk assessments. This should be kept in your study file, which should be readily available for audit purposes. You will be given a two week notice period if your project is to be audited. There is a checklist listing examples of documents to be kept which is available at http://ris.leeds.ac.uk/EthicsAudits.

We welcome feedback on your experience of the ethical review process and suggestions for improvement. Please email any comments to FMHUniEthics@leeds.ac.uk.

I hope the study continues to go well.

Best wishes
Rachel

Rachel de Souza, Research Ethics & Governance Administrator, The Secretariat, Room 9.29, Level 9, Worsley Building, Clarendon Way, University of Leeds, LS2 9NL, Tel: 0113 3431642, r.e.desouza@leeds.ac.uk

Appendix G: Approval from University of York

UNIVERSITY of York

The Department of Health Sciences

Faculty of Science
York Trials Unit
Ground Floor, ARRC Building
University of York
Heslington, York YO10 5DD

Direct line (01904) 321374
Email: catherine.hewitt@york.ac.uk

www.york.ac.uk/healthsciences

Our ref: DR_UKATT_01

Date: 14/05/19

Dear Sir/Madam,

York Trials Unit received a request to access the UK Alcohol Treatment Trial (UKATT) data listed below on 17/12/2018 from 'Kate Thresher' to investigate the psychometric properties of the Brief Addiction Therapist Scale (BATS):

- Alcohol consumption (Form 90): Pre-treatment assessment, three-month and one-year follow-up data from the participants in Motivational Enhancement Therapy (MET) intervention group.
- Sociodemographic information of the participants in Motivational Enhancement Therapy (MET) intervention group.
- Sociodemographic information of therapists in Motivational Enhancement Therapy (MET) intervention group.

There were three other individuals named on the data request (Bridgette Bewick, Gary Latchford and Helen Crosby) alongside the above named applicant. I can confirm that we received chief investigator approval from Gillain Tober for the release of the requested data on '06/11/2018'.

Yours faithfully

Catherine Hewitt

Deputy Director York Trials Unit and Professor of Trials and Statistic

Appendix H: Patient characteristics across therapist

Table H: Patient characteristics across the thirteen therapists

Therapist	Mean age pts	Sex	Ethnicity	Employment Status %	% Receiving benefits	Qual (% with)	Marital status (%)	Average no. of people in household	Housing (Owner occupied vs Rented)	Children (% with)	Mean DDD data at baseline
1	44.51	F=6 M= 7	12 white, 1 black Caribbean	46% employed, 23% sick, 23% unemployed	38%	85%	Mrd - 62%, S - 38%	2	46% owner, 54% rented	23%	17.79
2	38.87	F = 7 M=24	29 white, 2 black other	48% employed, 26% sick, 29% unemployed	48%	74%	Mrd - 55% S - 45%	3	45% owner, 55% rented	52%	20.24
3	37.75	F = 2 M = 11	13 white	38% employed, 23% sick, 38% unemployed	54%	61%	Mrd - 23% S- 77%	2	31% owner, 69% rented	15%	23.85
4	40.1	F = 3 M = 14	1 black 2 asian 14 white	35% employed, 41% sick, 24% unemployed	53%	59%	Mrd - 65% S - 35%	3	35% owner, 65% rented	58%	24.54
5	44.07	F = 9 M = 20	29 white	21% employed, 3% retired, 52% sick, 24% unemployed	52%	62%	Mrd - 45% S - 55%	3	62% owner, 38% rented	34%	29.78

6	38.58	F = 3 M= 28	31 white	23% employed, 58% sick, 16% unemployed	55%	58%	Mrd - 29% S - 70%	5	19% owner, 81% rented	23%	28.46
7	42.78	F = 15 M = 26	31 white	44% employed, 29% sick, 2% retired, 2% student, 2% volunteer, 20% unemployed	41%	32%	Mrd - 56% S - 44%	3	27% owner, 73% rented	24%	19.57
8	42.82	F = 9 M = 30	39 white	44% employed, 33% sick, 15% unemployed, 8 retired	56%	66%	Mrd - 31% S - 69%	2	21% owner, 29% rent ed	21%	27.31
9	38.64	F = 7 M=23	1 asian 29 white	40% employed, 33% sick, 27% unemployed	43%	73%	Mrd - 43% S - 57%	3	33% owner, 67% rented	43%	26.67
10	42.81	F = 10 M = 22	32 white	28% employed, 34% sick, 31% unemployed, 6% retired	38%	56%	Mrd - 31% S - 66%	2	44% owner, 56% rent	16%	22.90
11	41.10	F= 8 M= 17	25 white	40% employed, 32% sick, 12% retired, 12% unemployed, 4% volunteer	28%	64%	Mrd - 52% S - 48%	3	40% owner, 60% rental	36%	22.64

12	42.54	F = 3 M -= 10	13 white	38% employed, 15% sick, 15% retired, 23% unemployed	38%	69%	Mrd - 30% S - 69%	2	30% owner, 69% rented	69%	22.83
13	43.81	M= 7	7 white	43% employed 29% retired, 29% sick	0%	57%	Mrd - 71% S - 29%	4	71% owner, 29% rented	71%	19.45

F= Female; M= Male; Mrd = Married; Pts = Patients; Qual = Qualifications; S = Single