

Copyright is owned by the Author of the thesis. Permission is given for a copy to be downloaded by an individual for the purpose of research and private study only. The thesis may not be reproduced elsewhere without the permission of the Author.

Running head: EXPECTED CHANGE IN CBT

**THE EXPECTED COURSE OF CHANGE FOR CLIENTS
UNDERTAKING COGNITIVE BEHAVIOURAL THERAPY AS PREDICTED BY
EXPERIENCED AND NOVICE CLINICIANS**

A thesis presented in partial fulfilment of the requirements for the degree of:

Doctor of Clinical Psychology

Massey University

Palmerston North, New Zealand

Amber Fletcher

February, 2015

ACKNOWLEDGEMENTS

The completion of a doctoral degree requires commitment, strength and determination, not only from within, but from the people a candidate surrounds oneself with. The people that I list below have made navigating through challenges much more manageable and celebrating successes far more meaningful throughout my doctoral process.

First and foremost, I could not have completed my doctoral thesis without the support of my supervisors, Professor Emeritus Ian Evans, Cheryl Woolley, and Dr Steve Humphries. Ian, there is no doubt that I owe you many thanks for improving my abilities as an academic writer, but I also thank you for your continued support, patience, and belief in my capabilities as a researcher. Cheryl, I thank you for also being patient as well as for your ongoing encouragement, attention to detail, and dedication in supporting me to complete both my thesis and my internship. Steve, the sincere gratitude I have for you cannot be expressed in words but I will be forever grateful to you for enduring the constant emails, numerous queries, and days that I am sure you shared my frustrations (and almost tearing out of hair). Thank you for your knowledge, your patience, your calmness, and your kindness.

Thanks are also offered to Malcolm Loudon for going beyond our expectations in realising our electronic task – it was instrumental in the success of this thesis. Thank you to Dr Gus Haberman for your support as a secondary supervisor to Steve, and for supporting me in succeeding with my confirmation process.

To my three biggest cheerleaders; my partner Phillip, my wonderful friend Amanda, and my mother, Joyce, thank you for the unwavering support, your attempts to understand my frustrations, the many treats, and for reminding me that “working hard is important, but there is something that matters more, believing in yourself.” This all would have been unfeasible without you. I also extend my thanks to Yayo, my sisters Chrissie and Sonja, my stepfather Rob, Nana and Grandad, my beautiful nieces and nephew, Angel, Ryden, and Rubi-Rose, Andy and Diane, Livy, Veronica, Emma, Jared, Sam V, Sam R and Rebekah, Sarah and all my other friends and family that have provided me with support as well as multiple moments of joy that were very much needed.

For my clinical training in general I would like to thank my intern buddies, Rachael and Karen. I could not have wished for better support throughout the clinical process, and enjoyed knowing that there was someone there who completely understood what I was going through. Rachael in particular, I am so grateful to the programme for bringing your friendship into my life. Liz, I owe you my gratitude for your support early on in my clinical study and for your continued support throughout. Edwin, my office buddy for a short period and person to bounce my thoughts off, thank you for providing me with laughs over lame jokes, sharing your sweet treats and just offering a listening ear throughout the late stages of my internship and thesis. A huge thank you also goes to Jan Dickson for putting all her effort into reducing as much stress as possible throughout our training process; you are a beautiful, wonderful person. My clinical supervisor, Dr Erin Mooney, words cannot express how much I owe you for the wisdom, knowledge, and encouragement you still continue to provide me and for inspiring me to be the best clinician I can be. To the CAFS staff and other current and post clinical students, thank you for always motivating me, being compassionate, and validating my frustrations with challenges I faced.

ABSTRACT

Change, in the direction of improvement, is one of the main outcomes sought when treating mental health issues such as depression and anxiety. Historically, the focus of research has centred on change following the end of therapy, with recent discussions indicating that to promote better practice, understanding how the individual client changes, session-by-session, over the course of therapy is paramount. By incorporating a measure of progress at each session, it is proposed that clinicians will improve their ability to determine what reflects progress for clients, when intervention is required, and which aspects of therapy must be prioritised. Furthermore, the scientist-practitioner gap, where deficiencies in how practice influences research and how research influences practise have been identified and may be managed by actively collecting data about client progress. Practicing clinicians can then utilise research methods to both understand their own practice as well as provide insight into their practice that can influence further investigations. Using the primary therapy modality used to train New Zealand clinical psychologists, cognitive behavioural therapy (CBT), the present study aimed to provide an insight into what pattern of change was expected by both experienced and trainee clinicians when considering a client with depression and a client with anxiety undertaking a 12 session protocol of CBT, and how this compared to the current research literature. In addition, this study aimed to identify the pattern of change that was expected to occur for each client when considering overall symptoms, mood, and behavioural change. This was done by inviting experienced and trainee clinicians to complete an online task/questionnaire where participants were encouraged to plot session-by-session scores on three separate measures pertaining to each type of change using a specially designed graph. Despite the limitations of using hypothetical cases, findings showed that there

were no significant differences in predictions made by experienced or trainee clinicians, with clinicians overall predicting a decelerating curvilinear progression of change. When explored further, results indicated that clinician predictions differed from the research literature in a number of ways. Whether or not this can be attributed to lack of awareness of the research literature, or is reflective of the true nature of clinical practice, still requires further exploration.

TABLE OF CONTENTS

ACKNOWLEDGEMENTS	ii
ABSTRACT.....	v
TABLE OF CONTENTS.....	vii
LIST OF FIGURES	ix
LIST OF TABLES	xi
CHAPTER ONE: INTRODUCTION	1
CHAPTER TWO: PSYCHOTHERAPY OUTCOME RESEARCH.....	3
Classic Outcome Studies: Efficacy, Effectiveness, and Process-outcome Studies	4
What Do We Mean by a Good Outcome?	22
CHAPTER THREE: PERSPECTIVES ON CHANGE	33
Theories and Principles of Change	33
Other Aspects Related to the Pattern of Change.....	50
Client Factors and Change	58
Cognitive Behavioural Therapy and Change	63
Interrelationships Between Overall Symptoms, Mood, and Behaviour Change	65
Summary	70
CHAPTER FOUR: THE NATURE OF THE CURRENT RESEARCH	73
CHAPTER FIVE: METHOD.....	79
Ethics.....	79
Methodological Framework.....	79
Participants.....	81
Measures	82
Data Collection	94
Analyses	95
CHAPTER SIX: RESULTS	121
Demographics	121
Preliminary Analyses: Assumption Testing.....	122
Analysis 1: Clinician Awareness of Change Theories.....	124
Mr T	125
Analysis 2: Shape of Change over Sessions 2 -12.....	125
Analysis 3: Presence of Discontinuous Patterns of Change	145
Analysis 4: Continuous Change versus Discontinuous Change	146
Analysis 5: The Relationship between Overall Symptom, Mood and Behavioural Change ..	147
Analysis 6: Clinically Significant Symptom Reduction by the Final Session of Therapy	148

Analysis 7: Maintenance of Therapeutic Change at 3- and 6-Month Follow-up	149
Ms S	154
Analysis 2: Shape of Change over Session 2 -12	154
Analysis 3: Presence of Discontinuous Patterns of Change	173
Analysis 4: Gradual, Steady Change versus Variable Change	173
Analysis 5: The Relationship between Overall Symptom, Mood, and Behavioural Change ..	174
Analysis 6: Clinically significant symptom reduction by the final session of therapy.....	175
Analysis 7: Maintenance of therapeutic change at 3- and 6-month follow-up.....	176
CHAPTER SEVEN: DISCUSSION	183
Summary of Key Findings	183
Clinical and Practical Implications	193
Limitations	196
Conclusion	198
REFERENCES	201
APPENDIX A: ELECTRONIC TASK	238

LIST OF FIGURES

Figure 3.1. Example of an early rapid response using Beck Depression Inventory scores.
53

Figure 3.2. Example of a sudden gain using Beck Depression Inventory scores.....56

Figure 3.3. Example of a depression spike using Beck Depression Inventory scores.58

Figure 5.1. Examples of linear, quadratic, and cubic curves..... 108

Figure 5.2. Example of predictions reflecting a smooth and continuous predicted trend.
 111

Figure 5.3. Example of predictions reflecting a discontinuous predicted trend..... 111

Figure 5.4. Original graph and graph adjusted to show scores progressing in the same
 direction for client with depression..... 113

Figure 5.5. Original graph and graph adjusted to show scores progressing in the same
 direction for client with anxiety..... 114

Figure 5.6. Pre- and post-test scores and cut-off a, cut-off b, and cut-off c indications for
 participant x (Jacobson & Truax, 1991; p14). 117

Figure 6.1. Residual plot for negative affect score (PANAS-NA scores)..... 122

Figure 6.2. Standardised residual scatter plot for negative affect across time (PANAS-NA
 change)..... 123

Figure 6.3. Average BDI-II, PANAS-NA and No. of Activities score for Session 12, 3
 month follow-up and 6 month follow-up..... 150

Figure 6.4. Average BAI, PSWQ and percentage of activities score for Session 12, 3 month follow-up and 6 month follow-up.....177

LIST OF TABLES

Table 6.1. <i>Results of Fitting a Taxonomy of Multilevel Models for Change on the BDI-II Across the Course of Therapy</i>	126
Table 6.2. <i>Results of Fitting a Taxonomy of Multilevel Models for Change on the PANAS-NA Across the Course of Therapy</i>	134
Table 6.3. <i>Results of Fitting a Taxonomy of Multilevel Models for Change in the Number of Activities Participated across the Course of Therapy</i>	140
Table 6.4. <i>Frequency of Each Sequence of Measured Dimension Expected for a Client with Depression</i>	147
Table 6.5. <i>Significance of Mauchly's Test of Sphericity for Session 12, 3 Month Follow-up and 6 Month Follow-up for Overall Symptom Change (BDI-II), Negative Affect (PANAS-NA) and Number of Activities</i>	149
Table 6.6. <i>Results of Fitting a Taxonomy of Multilevel Models for Change on the BAI Across the Course of Therapy</i>	156
Table 6.7. <i>Results of Fitting a Taxonomy of Multilevel Models for Change on the PSWQ Across the Course of Therapy</i>	164
Table 6.8. <i>Results of Fitting a Taxonomy of Multilevel Models for Percent of Activities Participated In over the Course of Therapy</i>	168
Table 6.9. <i>Frequency of Each Sequence of Measured Dimension Expected for a Client with Anxiety</i>	175

Table 6.10. <i>Significance of Mauchly's Test of Sphericity for Session 12, 3 Month Follow-up and 6 Month Follow-up for Overall Symptom Change (BAI), Level of Worry (PSWQ) and Percentage of Activities Participated In</i>	176
--	-----

CHAPTER ONE:INTRODUCTION

This thesis is about the expectations and judgments of clinical psychologists, both novices and experts, when asked to consider the sorts of changes they would anticipate in clients undergoing therapy for common clinical disorders. It is an *analogue* study in the sense that the participants were asked to judge likely progress in two hypothetical cases described to them. They were asked to make predictions of client status after each session in order to see if they expected change to be a simple, steady, linear improvement as predicted by a number of theories of change in different psychological dimensions; behaviour, mood, and clinical symptoms, or whether some other pattern was predicted.

My interest in understanding change is based on my desire to become a clinical psychologist since I began high school, and the training that has led to the realization of this goal. As a New Zealand Doctor of Clinical Psychology candidate, I have been trained with Cognitive Behavioural Therapy (CBT) as my guiding treatment modality. As such, and based on CBT being empirically supported in the treatment of numerous mental disorders, understanding change while undergoing this form of treatment is the focus for the current study. Like any clinician, as I have progressed through my training my interest in how clients change has deepened, based on the aspiration to support my clients as effectively and efficiently as possible. This then motivated me into thinking of ways to better understand the process of change by particularly considering a treatment modality highly supported by the empirical research literature.

More often than not the research literature discusses therapeutic change in terms of before and after measurements of treatment outcome, but there is a much broader and longer tradition of measuring change within sessions or during the course of treatment, for example, the single case design in Applied Behaviour Analysis (Barlow, Hayes, &

Nelson-Gray, 1984). Despite this way of understanding change being present for a long period of time, only recently has research shifted its emphasis from the overall outcome of treatment to the measurement of change across the course of therapy (Trauer, 2010c).

There is some literature, to be reviewed in the following chapters, which suggests that there are various possible patterns of change, and the purpose of the study was to examine the degree to which clinical practitioners would foresee such patterns, or expect others. In order to provide an introduction to the research, therefore, the following chapter will selectively review (a) the literature on the importance of outcome measurement in psychotherapy research, including session by session progress, (b) the different ways of measuring outcomes that have been reported during and after therapy, and (c) some theories highlighting change patterns that suggest therapeutic progress does not simply follow a linear pathway.

The relevance of this study to practice is that by gaining an understanding of clinicians' expectancies, and the extent to which they fit empirical findings and theoretical models about change, it should enable clinicians to be alerted to possible subtle changes in clients during the course of therapy. Practitioners would be able to use this information to maximise individual client outcomes as well as better judge the merits of empirical studies supposedly validating different techniques in CBT.

CHAPTER TWO: PSYCHOTHERAPY OUTCOME RESEARCH

In the practice of clinical psychology, one of the main goals is to work alongside clients in achieving a reduction in distress and impairment in their daily life and a return to, and maintenance of, normal functioning in everyday life (Lambert & Hill, 1994). Following the inception of psychotherapy over a century ago, its ability to achieve such successful outcomes has been under empirical scrutiny. Interest, and doubt, was further stimulated by Eysenck's (1952) well-renowned paper claiming there was no evidence that psychotherapy was more effective than no therapy at all. This inspired researchers to direct their attention to providing empirical support to refute this claim, resulting in hundreds of studies that have shown psychotherapy, in different forms and treating various disorders, to be potentially beneficial (Consumer Reports, 1995; Kopta, Lueger, Saunders, & Howard, 1999).

Since the emergence of psychotherapy, an estimated 500 different approaches have been made available (Kazdin, 2007). Therefore, to ensure that clients are being provided with a therapy that is likely to be beneficial for their presenting problems, it is crucial that practicing clinicians use treatments that have empirical support (Kazdin, 2007). The first attempt to establish recommendations based on empirical support for therapies was undertaken 25 years ago, through the Quality Assurance Project (1982; Quality Assurance Project, 1983; Quality Assurance Project, 1984; Quality Assurance Project, 1985a; Quality Assurance Project, 1985b). Sixteen years ago, Division 12 of the American Psychological Association (APA, Chambless & Hollon, 1998) undertook an exercise for creating recommendations for treating particular disorders with therapies based on the empirical validation from efficacy, effectiveness and process-outcome studies.

Classic Outcome Studies: Efficacy, Effectiveness, and Process-outcome Studies

Efficacy studies involve the comparison of therapies treating a specific, singular disorder with placebo or alternative therapies. Usually volunteer participants meeting stringent inclusion and exclusion criteria are randomly assigned to treatment or control groups. Such randomized control is considered by many researchers to be the gold standard for treatment evaluation. The therapies are administered in a highly structured manner and completed in a fixed number of sessions, with clients typically assigned to each condition randomly. The aim of this high methodological control is to strengthen internal validity, and establish cause-and-effect inferences (Howard, Moras, Brill, Martinovich, & Lutz, 1996; Leon, Kopta, Howard, & Lutz, 1999).

Although efficacy studies provide good quality information about which therapies have demonstrated success in treating certain syndromes, there can be a vast difference between what works in actual clinical practice as opposed to highly controlled laboratory settings. This has left some theorising that results provided by efficacy research are perhaps difficult to generalise to clinical practice (Seligman, 1995). Seligman (1995) noted a number of limitations with the inflexibility of efficacy research including how therapy is rarely allocated a fixed duration prior to the first meeting, unless controlled by funding. Even when a clinician is following a manual, extra sessions may be required to address unexpected issues that arise. In practice, psychotherapy is typically self-correcting, and a technique is discontinued if a client is not responding to it. With a large number of clients often entering therapy with more than one mental health issue (i.e., more than one diagnosis) another problem arises due to participants in efficacy studies only being included if they can be classified as experiencing a single mental health

diagnosis. Thus, efficacy studies do not appear to reflect the true nature of clients with mental health difficulties. Seligman (1995) also noted that these studies focus on the cessation of the symptoms of a specific problem, rather than the more subjective improvement of general functioning.

Effectiveness research, on the other hand, was developed based on the aforementioned issue with the generalisability of efficacy studies (Borkovec, Echemendia, Ragusea, & Ruiz, 2001). It thus extends the information provided by efficacy studies by demonstrating whether or not therapies work in the clinical setting, doing this by evaluating clients who are already in therapy (Howard et al., 1996), or conducting trials with real clients who are presenting at clinics often with complex concerns. Despite overcoming the limitations of efficacy studies being the purpose of effectiveness studies, this form of research has its own criticisms. While it is an advantage that therapies used in effectiveness studies are more flexible, it is felt that this method does not account for individual characteristics, thus impacting on each client's rate and extent of change, an area of clinical interest (Lutz, Lowry, Kopta, Einstein, & Howard, 2001). Furthermore, information in effectiveness studies may not provide a comprehensive picture of who is benefiting, or not benefiting, from the therapies studied. This is thought to be due to clients having to agree to participate in research after already initiating contact with a mental health service delivery system, with those who consent possibly having a higher likelihood of achieving a positive outcome based on their commitment and motivation to participate in therapy (Lutz et al., 2001).

While there are issues specific to their respective methodologies, a problem faced by both efficacy and effectiveness studies is that findings fail to answer the question "Is this particular therapy working for this particular client?" (Lutz, Martinovich, & Howard,

1999). A research area providing further understanding of client change for the past 50 years (Orlinsky, Ronnestad, & Willutzki, 2004), termed process-outcome research (Timulak, 2008), is one broad investigative strategy in which answers to this question have been sought. It aims to produce findings that can inform the therapeutic process in real time, focusing on three key areas. The first area, involving understanding the type of individual and the conditions in which psychotherapy elicits change, relates to both process and outcome. Termed moderators, these factors occur prior to the initial therapy session and improve the matching of clients to therapy and techniques (Kazdin, 2007). Examples of moderators of therapeutic change include age, sex, ethnicity, temperament, or age cohort (e.g., baby boomers) (Kazdin, 2007). In outcome research, these variables are assessed for their impacts on the direction or magnitude of the association between the intervention and outcome (Kazdin, 2007).

Secondly, process and outcome research addresses factors that allow change to occur, termed mediators and mechanisms of change (Newnham & Page, 2010). Mediators of change are intervening variables that have been found to have a potential statistical association between the independent and dependent variable, although the causal processes involved in creating the pathway may be unknown. Research on mechanisms focuses on the steps and processes that produce change. Researchers have focused on understanding what is essential in treatment to produce change, and the way in which this change is achieved. Without such information the clinical community only knows that change occurs without knowing how it occurs (Evans, 2013; Kazdin, 2007). Furthermore, it is important to identify what variable is responsible for which effect, for example, whether or not therapeutic alliance reduces symptoms, or symptom reduction improves the therapeutic alliance. By doing so, it is the hope of researchers to enable clinicians to more effectively assist their clients (Lambert & Hill, 1994).

The final area of process-outcome research concerns the shape of change during the course of therapy (Kazdin, 2007; Laurenceau, Hayes, & Feldman, 2007). With the majority of research focusing on pre- and post-treatment comparisons (Groth-Marnat, 2003), more recent research is now directed at understanding the evolving and underlying trajectory that change progresses through between these two points. This can be done by using outcome measures session-by-session or frequently enough to reflect change over the duration of psychotherapy (Lambert, 2010).

Likewise, with efficacy and effectiveness studies, process-outcome research has been found to have limitations, with this area of research still largely conducted in the research setting, rather than through actual clinical practice. As suggested by Weisz and colleagues (Weisz, Donenberg, Han, & Weiss, 1995), the validity of findings will be increased by understanding process and outcome within a naturalistic setting as opposed to controlled lab conditions. Process and outcome research also incurs problems with psychotherapy research having a significant focus on the process of therapy, but minimal interest in understanding mediators and mechanisms (Evans, 2013; Kazdin, 2007). Additional issues arise with theorists acknowledging that clients with almost identical characteristics will not always follow identical pathways through therapy.

Overall, results derived from the three branches of classic outcome study methodologies show consistencies in certain therapies, techniques, and concepts, expressing a complex but comprehensible view of therapy that highlights both common factors across therapies and specific therapeutic concepts that are important in terms of outcome (Orlinsky et al., 2004). Whilst acknowledging the valuable information that efficacy, effectiveness and process-outcome studies provide, and despite attempts to address the limitations acknowledged in these classic outcome studies, one problem

persists: the implementation of the evidence base in clinical practice and contribution of clinical practice to the evidence base, the predicament otherwise known as the scientist-practitioner gap.

The scientist-practitioner gap

The practice of clinical psychology has long been based on principles of the scientist practitioner model, in which research influences practice and practice is influenced by research, continuing to do so in New Zealand, Australia, North America, and Britain (Evans & Fitzgerald, 2007). This requirement of therapy being guided by an evidence base is further emphasised in Core Competencies for psychologists, defined by the New Zealand Psychologists Board, indicating that psychologists have an obligation to use “scientific evidence to inform and guide practice (New Zealand Psychologists Board, 2011, p. 4)” and, as such, these guidelines are upheld by services offering psychotherapeutic support provided by psychologists such as through District Health Boards. It thus seems surprising that researchers and clinicians both acknowledge a large gap between research and practice (Newnham & Page, 2010). Whilst the professional training for researchers and practitioners is quite similar, the two groups appear to end up practicing in different worlds, and unfortunately, to some extent, have difficulty understanding the needs of each other and how they can collaborate to achieve a sound basis for psychotherapy (Goldfried & Wolfe, 1996; Wolfe, 2012).

From the clinicians’ perspective, problems with the implementation of outcomes derived from randomised controlled trials (RCTs), the “gold standard,” exist because clinicians feel that researchers focus on diagnosed disorders more so than the clients and, as such, do not adhere to the complexities that are present in individuals who attend psychotherapy (Wolfe, 2012). This includes clients who present with more than one

problem area, the need for clinicians to attend to problems which they or the client feel are more important, and problems that impact on the clients' difficulties that are not identified as the clinical problem (Stiles, Honos-Webb, & Surko, 1998; Wolfe, 2012). Again, this is indicative of the limitation problems suggested, regarding efficacy and effectiveness studies, leaving clinicians to prefer to use prior experience, seek supervision or review with colleagues to guide and direct their clinical practice (Eatwell & Wilson, 2007).

From the researcher perspective, it is felt that the RCT method provides valuable and incontestable information, whereas clinical judgement and experience are less valid (Hamilton, Kendall, Gosch, Furr, & Sood, 2008; Wolfe, 2012). Researchers also feel that the study of psychotherapy allows for flexibility and thus does, in some ways, account for individual differences (Wolfe, 2012).

The scientist-practitioner gap has not been evaluated in New Zealand. However, based on the contention within other Western countries (i.e. UK, USA from where New Zealand derives a large amount of information), it is an area that requires further clarification. In addition to New Zealand clinical information being derived from other countries, the resulting knowledge may not be relevant to the New Zealand population, especially in terms of New Zealand's indigenous population who make up a significant proportion of mental health consumers (Baxter, Kingi, Tapsell, Durie, & McGee, 2006; Evans & Fitzgerald, 2007). In addition to potential irrelevance, New Zealand clinicians may not have access to the resources required to implement research recommendations (Evans & Fitzgerald, 2007). While some would assume that the obvious remedy to this would be to develop home-grown research where there has definitely been progress, it is unlikely that it will ever reach the scale of other countries that have a higher

clinician/researcher proportion, bigger sample sizes and greater access to resources (Evans & Fitzgerald, 2007).

It is clear from the discussion on the scientist-practitioner gap that clinicians and researchers work in different worlds and value different findings. Whereas researchers largely appear content with group means and understanding the “average” client, clinicians are more focussed on the specific and concrete understanding of a therapy, and how and why it supports an individual in changing (Wolfe, 2012). I, personally, have found that in clinical practice there have been times where I have seen a client with a problem that is not accounted for by best-evidence practice, while also finding empirically validated treatments to lack effective results in clients for which it has been recommended. As both research and clinical practice aim to find the best ways to treat particular disorders in individuals, I find myself agreeing with Wolfe (2012) wherein both areas have benefits and limitations that, with an effort to promote the combination of both valuable methods to the evidence-base, is likely to improve the understanding of what works in therapy for whom.

With both sides of science and practice agreeing that the scientist-practitioner gap is undesirable, suggestions have been put forth of ways to address it. A special section was even dedicated to the cause in Issue 49 Volume 2 of *Psychotherapy*, developed by the American Psychology Associations (Division 12) 2011 Committee on Science and Practice (Teachman et al., 2012). Primarily, and acknowledging that the findings provided by research are valuable, it is recommended that supplemental information is required to make findings more applicable to clinical practice. One suggestion is the inclusion of case studies that provide different perspectives and interpretations of an empirically supported treatment when applied to the clinical setting (Barkham et al.,

1998; Wolfe, 2012). Further clinician involvement in research has also been encouraged, for example, Goldfried (2011) developed surveys for clinicians to indicate successes and problems with empirically supported treatments. In addition, the presentation of studies that show both good and poor outcome can reveal further information about what works and what does not in treatment studies (Wolfe, 2012). Such a view also supports the notion of moving toward identifying important principles as opposed to treatment modalities. Working towards an understanding of the underlying process of change can further promote research integration into practice by expanding the knowledge base on mediators and moderators of change (Wolfe, 2012).

Patient-focused research

Addressing the limitations identified in the classic outcome studies by taking into account how clients change in therapy, and developed with the view to reduce the scientist-practitioner gap, Howard and colleagues (Howard et al., 1996) developed the Patient-focused research method. This methodology involves routinely monitoring an individual's progress while they are involved in therapy, with the resulting data being provided as feedback to the clinician responsible. Patient-focused research was developed due to a growing use of reimbursement systems and specifications of accountability in the USA (APA Task Force on Evidence Based Practice, 2006) and the United Kingdom (Health and Care Professions Council, 2012) alongside growing desire for research that was easier to apply in the clinical setting (Nathan, Stuart, & Dolan, 2000). These ideas began to highlight that the outcome of therapy was being presented as the purchase for clients and not treatment itself (Lambert, 2010). To understand the benefits of treatment, it became important to understand the pathway to change rather than just whether or not treatment is effective (Lambert, 2010). In addition, Howard and colleagues (1996)

specified that the use of group means and associated inferential statistics were too general to provide results applicable to a single individual, thus invalidating the sole use of efficacy and effectiveness results (Krause, Howard, & Lutz, 1998; Lambert, 2010).

While the purpose of efficacy and effectiveness studies is to determine what types of therapies or techniques produce positive outcomes, patient-focused research is interested in identifying what factors need to be considered to know whether a treatment is appropriate for an individual at a particular point in time for a particular type of problem. This method also focuses on ways in which therapists can identify clients who are not achieving positive gains in a small amount of time, so as to avoid clients deteriorating or dropping out of therapy without any positive gains (Lambert, 2010). Additionally, changes different from the previous week can be addressed within session to understand what worked well and what needs to be altered (Hayes, Miller, Hope, Heimberg, & Juster, 2008). Other beneficial factors of this methodology include the ability to evaluate the expected effectiveness of treatment for a particular client prior to starting it.

Patient-focused methodology aims to overcome the barriers posed by the use of group means by emphasizing the use of individual repeated outcome assessment throughout the therapeutic process. Further extensions on previous methodologies include concentrating on “positive” outcomes as demonstrated by clinical significance, as opposed to a more generally and vaguely termed successful outcome, to provide findings applicable to those who do not respond to therapy or deteriorate (Lambert, 2010). As patient-focused research does not rely on group means, it can reveal not only individual differences but also notable consistencies within groups of patients with similar characteristics, which ultimately will extend our knowledge of what changes occur, and

for whom (Lambert, 2010). Consequently, this form of research allows clinicians to understand how clients who have similar characteristics progress through therapy by addressing their similarities and differences in responding (Lutz, 2003). Patient-focused research allows for the understanding of process and outcome by comparing treatments in terms of their “dose-response”, that is the number of sessions before positive, meaningful change has been achieved, as well as the final outcome (Lutz, 2003). Furthermore, this method enables the observation, verification, and accumulation of the experiences and practices of clinicians in their clinical settings, which enhances external validity and ultimately is helpful in improving the quality of therapy (Barlow et al., 1984; Lambert, Whipple, Smart, Vermeersch, & Nielsen, 2001b). While this research is indeed promising, there appears to be a lack of understanding of how current reliable and valid psychometrics can be implemented in a similar way (Goodman, McKay, & DePhilippis, 2013). Those currently involved in this research area have focused on using their own measure, the Outcome Questionnaire, to monitor progress in therapy by reviewing symptomatic functioning, interpersonal problems, and social role performance (discussed in the later section “Statistical Predictions and Feedback”).

Acknowledging that evaluating change processes in therapy is exceedingly complex due to variability across theories, the premise of evaluating change at each session can provide helpful overall information on an individual basis. This allows the separation and analysis of individual and therapeutic effects (Lambert & Hill, 1994) as well as creating a more precise index of when change is occurring (Hayes et al., 2008). This then enhances a clinician’s ability to use techniques that are better suited to the client which can further enhance treatment response (Newnham & Page, 2007). It also allows clinicians and researchers to know the effect of each session, and what variables this

corresponds to (Hansen & Lambert, 2003), further indicating the importance of the inclusion of frequently used psychometrics in this form of research.

With the main aim of enabling better practice, patient-focused research and its findings are thought to be of interest to six parties; clients, secondary purchasers, clinicians, managers, sponsors, and researchers. Most importantly, the clients receiving therapy need reassurance that they are being treated with an intervention shown to effectively improve individuals with similar psychological impairment, and to some extent, similar circumstances. Confidence regarding the treatment provided for clients is also important for secondary purchasers, those who do not directly receive treatment, such as parents, employers, and institutions. With the importance that is placed upon the scientist-practitioner gap, clinicians too need to possess the knowledge that the therapy they are delivering is appropriate and have the ability to achieve successful outcomes in the clients they treat. Represented by managers and sponsors, organisations such as private practices and District Health Boards need to be assured that they are allocating appropriate funds towards particular theoretical orientations shown to be successful in treating certain disorders. With the expansion of the current breadth of knowledge, and the creation of areas which require investigation, researchers also benefit from the patient-focused position. With the aim of understanding how a certain individual will respond to therapy, all six parties described will benefit.

Acknowledging that patient-focused research is advantageous in understanding the individual client and how they change, and the previously discussed limitations of efficacy and effectiveness studies in terms of clinical practice, this does not reduce the importance and necessity of such studies (Lambert, Hansen, & Finch, 2001a). While efficacy and effectiveness research alone may not reveal the true nature of change within

the therapeutic setting; by combining the information provided by these three forms of research, the issues surrounding the generalisation of efficacy and effectiveness studies to the clinical setting will be reduced (Newnham & Page, 2007).

Routine outcome measurement

One of the key focuses of patient-focused research is the use of routine outcome measurement. Routine outcome measurement, in which outcome measures are used periodically over the sessions in the therapeutic process, and dating back to methods used in early Applied Behavioural Analysis (Barlow et al., 1984), promotes communication between client and clinician by enhancing treatment planning and keeping track of which goals require attention. Progress is easily followed, allowing the client to visibly see their progress which can promote empowerment, self-efficacy, and recovery. Clinicians also benefit, with frequent outcome assessment enhancing the identification of issues not overtly acknowledged in therapy such as whether or not clients are responding appropriately to therapy (Trauer, 2010b). The frequent use of outcome measures improves knowledge, for both clinicians and researchers, in regards to which measure provides the best evaluation of certain mental health aspects, and their appropriateness in certain settings (Trauer, 2010b). Trainee clinicians may also benefit from routine outcome measurement as this can support their understanding of how clients change, increase reflection on clinical practice and how it affects clients, as well as build confidence. In addition to clients and clinicians, policymakers and those who fund mental health practices can increase their awareness of which treatment protocols are practical and beneficial through comparisons with other services available, and highlighting areas in which improvements can be made (Garland, Kruse, & Aarons, 2003). As such, routine

outcome measurement is felt to be best practice in psychotherapeutic treatments (APA Task Force on Evidence Based Practice, 2006; Trauer, 2010b).

New Zealand has also begun to stress the importance of routine outcome measurement (Mellsop & Smith, 2010), although not to the extent of session-by-session measurement. Establishing the Classifications and Outcomes Study (CAOS) (Gaines et al., 2003), The New Zealand Mental Health Casemix Classification and Outcomes Study aimed to create a national case-mix classification, where episodes of care were classified based on factors deemed to best predict the need for, and cost of, care, with an aim to increase the routine use of selected outcome measures in New Zealand's psychiatric services. Following on from this study, mental health practitioners working in New Zealand District Health Board's (i.e. hospitals and government funded agencies), routine measurement is performed using the Health of the Nation Outcome Scale focusing on review of a client's change progression at the initial meeting, every three months while in service, and at discharge (Wing et al., 1998). While improving the opportunities in the use of routine outcome measurement, issues arise in understanding client change due to the less frequent implementation of such measures. The measures are currently also in the developmental stage and continue to require time to establish reliability and validity. In addition, the measures are used across mental health disciplines and look at the reduction in symptoms of functioning, and do not necessarily focus on reduction in mood, diagnosis specific symptoms, or behavioural change (Wing et al., 1998). These issues could suggest that, especially in terms of the practice of psychotherapy, already established measures used in clinical practice may provide another opportunity to explore the use of routine outcome measurement, while also providing valid information in regards to how clients change.

While deemed best practice, routine outcome measurement using the tools specific to the profession of clinical psychology, psychometrics (Eatwell & Wilson, 2007), are rarely the primary method used in monitoring client change in therapy (Trauer, 2010b). Instead, monitoring treatment response relies heavily on the clinician's subjective judgement and intuition, and client self-report (Hannan et al., 2005). When looking at the use of outcome measures in clinical practice, a study of 15,918 licensed American psychologists (Phelps, Eisman, & Kohout, 1998) showed that 29% had used some form of outcome assessment in their routine practice, with a similar percentage (37%) identified in a later study (Hatfield & Ogles, 2004). In addition to the earlier study, those who used outcome measures were more likely (40%) to be found in a medical setting rather than practicing in private (24%). When addressing intent, approximately half of participants in one study of 50 general mental health practitioners stated that they thought it was possible to include outcome measures in their routine clinical practice, with the other half stating it was not (Garland et al., 2003). These findings were further replicated in a study focusing on the treatment of adolescents (Bickman, Rosof-Williams, Salzer, & Summerfelt, 2000). Reflecting on these results, it would seem that despite the benefits that outcome measures promote, regular use of such measures to track change is not a significant concern for clinicians, and is unlikely to be unless mandated (Garland et al., 2003).

Although research in this area needs to be updated, Patchett-Andersen (1997) found that the use of outcome measures in New Zealand clinical practice was comparable to that of the aforementioned American research. In a survey of clinicians practicing in New Zealand, findings indicated that the clinical interview, the Beck Depression Inventory, the Beck Anxiety Inventory, and the State Trait Anxiety Inventory were the most utilised measures in assessing depression and anxiety (Patchett-Andersen, 1997).

Numerous reasons have been identified for why clinicians find it difficult to monitor client progress in their routine practice. These include the view that outcome measurement requires too much effort, is potentially intrusive for clients (Hatfield & Ogles, 2004; Lueger et al., 2001), could prompt unnecessary stress for the client (Garland et al., 2003), as well as being time-consuming (Garland et al., 2003; Hatfield & Ogles, 2004). Not solely restricted to the difficulties regarding administration and scoring, it was thought that routine outcome measurement enforced an expectation that clinicians spend more time proving what they do, rather than doing what they do. An example of this follows Evans (2005) attempts to support clinical psychology students to implement a personalised progress measure each week with their clients to monitor change progression focussing on themes such as the client's view of themselves, behavioural change, and changes in their world perspective. While the importance of change monitoring over the course of therapy was emphasised to the students, interest in utilising the measures was minimal (Evans, 2005).

One therapeutic modality that promotes routine outcome measurement, has received significant support through evidence derived from classic outcome measurements, and is one of the most utilised modalities in New Zealand clinical practice, is Cognitive Behavioural Therapy.

Cognitive behavioural therapy

Cognitive Behavioural Therapy (CBT) is a short-term and structured psychotherapy which has been generalised, and shown to be effective, to a number of disorders, including mood and anxiety disorders (Beck, 2011). To put it briefly, CBT supports individuals to overcome their current difficulties by focussing on unhelpful thinking and behaviour which results in emotional and behavioural change (Beck, 2011).

This is based on Beck's belief that dysfunctional thinking influences mood and behaviour, for example, if an individual has the thought "I cannot do anything," they may feel sad and opt to spend the remainder of the day in bed avoiding the completion of tasks that may have improved their mood. Therapists using CBT thus support a client in approaching their thinking styles in a more realistic and adaptive way (Beck, 2011). Such a way of approaching depression and anxiety has been shown to be effective with reviews of efficacy and effectiveness research suggesting that, despite clear indications of publication bias, CBT demonstrates superiority to no treatment or treatment as usual (Cuijpers et al., 2013; Hollon & Beck, 2013). Evidence also shows that CBT appears superior to psychopharmacotherapy alone and is as effective as other psychotherapeutic treatments (Hollon & Beck, 2013).

In understanding the specific techniques that result in client change in CBT, process-outcome studies remain limited, with limitations including a control or comparison group being absent, small sample sizes, and failing to assess mediators before symptoms (Crits-Christoph, Gibbons, & Mukherjee, 2013). As such, process-outcome research regarding CBT is mixed and requires more comprehensive research attention. Overall, emerging research supports the notion that clients typically show change reflective of a causal model, that is, CBT begins by targeting negative automatic thoughts which cause and maintain the client's psychological impairment, resulting in a reduction of depressive symptoms and improvement in mood alongside participation in positive behaviours (Oei, Bullbeck, & Campbell, 2006).

In terms of the New Zealand context, those who participate in study to become a clinical psychologist through a New Zealand university trained using CBT as their guiding therapeutic approach (Evans, 2002; Kazantzis & Deane, 1998). Perhaps due to

this emphasis, CBT has been shown to be the psychological approach more commonly used in clinical practice in New Zealand, with New Zealand having a greater proportion of therapists who use CBT as their principal approach compared to British and American psychologists (Kazantzis & Deane, 1998).

Similar to the emerging emphasis in outcome research, CBT views routine outcome measurement as best practice. This is based on CBT's goal-oriented nature, with textbooks often indicating that measurement of aspects of change, such as mood and behaviour, should be completed regularly throughout the course of therapy to monitor such goals (Beck, 2011; Clark, Hollifield, Leahy, & Beck, 2009; Westbrook, Kennerley, & Kirk, 2011).

CBT's empirical backing, familiarity with clinicians particularly in New Zealand, and its emphasis on routine outcome measurement has made it a prime candidate for evaluation and use in studies attempting to address the suggestions made from both classic outcome studies and commentators on the scientist practitioner gap. CBT furthermore has a clear idea of what should change during psychotherapy in terms of cognitive change resulting in mood and behaviour change. This provides a good basis for what to look for when clients change in CBT.

Summary

Ever since the development of psychotherapy, multiple efforts have been made to identify its worth. Efficacy and effectiveness studies alongside process-outcome research have provided numerous ways in which this has been done, and have successfully provided support for certain therapies in treating certain diagnoses. While helpful information is provided, the methodologies used create limitations when applying the evidence-base to clinical practice. This, in turn, has perpetuated the scientist practitioner

gap that has been observed in American and British clinicians. Recommendations concerning both areas indicate that when and how clients change will provide better information both to researchers and clinicians, with routine outcome measurement being encouraged to provide one opportunity of bridging that gap by engaging clinicians in research and researchers understanding clinician's needs. Emphasising the importance of routine outcome measurement opens up opportunities for using such predictions in real-life practice, which would then re-engage researchers and clinicians with the scientist-practitioner model. As the literature indicates that attempts to increase implementation of routine outcome measurement involves measures still in their developmental stages, the current study also aims to reflect on the use of current psychometrics that are already familiar to clinicians, in understanding how clients change. Furthermore, the issue of what outcomes are most important to measure and for how long they should last has become a major question for treatment research and is one of the themes guiding the present research: what do clinicians expect in terms of client progress, and in what domains of functioning? CBT, with its emphasis on routine outcome measurement, strong evidence-based backing, familiarity in New Zealand practice and theorised areas of change, has thus been identified as a modality of interest in terms of understanding how and why clients change. As such, CBT is the modality used in this study and commentary on certain domains will reflect on how this is viewed from a CBT perspective to ensure the study is placed in the appropriate context.

Having identified the limitations and difficulties experienced in the classic outcome literature, as well as the difficulties between researchers and clinicians, I will now focus on what is meant by a good outcome. As emerging research suggests how clients change across therapy is becoming of more importance than overall outcome, it is important to understand what is meant by change, and how it can be identified.

What Do We Mean by a Good Outcome?

In his book on how and why people change, Evans (2013) drew four distinct features of change from the literature. Firstly, there must be something needing to be changed, either by the individual or society. Secondly, literal change must occur, that is, an individual moves from one state of being to another. Thirdly, the end-point of change must meet some form of success, whether that is based on a social criterion, personal satisfaction, or both. Finally, change is achieved through a pattern or sequence. To provide an understanding of how good outcome can be understood, the following section discusses each of these features.

Knowing what to change

In research, it is no surprise that the goals are guided by the research question and thus the focus is typically on overall symptoms, as opposed to the symptoms broken down into separate areas such as mood or behaviour. In clinical practice, however, choosing the problematic behaviour requiring change is typically guided by the presenting problems and the client's goals for therapy. In CBT, the use of case formulation further highlights problematic areas that are maintaining a client's difficulties (Beck, 2011; Westbrook et al., 2011). If meeting for a mental health disorder, clinicians can be guided by the Diagnostic and Statistical Manual, now in its fifth edition (DSM-5, American Psychiatric Association, 2013), to identify the problem areas that typically align with a particular mental health difficulty. It is imperative to acknowledge, however, that clinical judgement is an important component in identifying difficulties that do not fit into the DSM's criteria. Coming to an understanding of what to change, the second feature of change concerns the occurrence of literal change where an individual moves from one state of being to another.

Literal change

When measuring therapeutic change in research, the focus is typically on symptom change through the use of self-report measures (such as the Beck Depression Inventory or the Symptom Checklist) and statistical significance, a mathematical equation used to identify significant change. Classic outcome studies have long used statistical significance to show genuine change has occurred. However, mounting criticism has suggested that such a method is not clinically relevant (Lambert & Hill, 1994). This is because statistical significance is thought to largely exclude the analysis of individual variability in terms of how specific individuals react to therapy. For example, due to the use of grouped subject means, results are unable to demonstrate how many participants are shown to benefit from therapy and how many are not (Jacobson, Follette, & Revenstorf, 1984; Kazdin, 1999; Lambert & Hill, 1994). Furthermore, a number of researchers maintain that statistical significance lacks meaningfulness to the client and clinical utility based on this analytical strategy's inability to show clinical change. For example, a small, trivial change can be deemed as statistically significant solely because it was achieved in a large sample which in turn has a substantial amount of statistical power to identify small changes (Hansen & Lambert, 2003; Hansen, Lambert, & Forman, 2002). In addition, it is important to acknowledge that just because change has been demonstrated at a statistical level, does not mean that the client feels that they have achieved meaningful change for themselves.

In trying to adapt a statistical method that understands clinically meaningful changes, clinical significance was developed, addressing Evans (2013) third feature of change where the end-point of change must meet some criterion of success.

Meaningful change

Jacobson and Truax's (1991) clinical significance method, a statistical method proposed to identify change that would have a noticeable effect on a client's life, has been consistently used in psychotherapy research, with the *Journal of Consulting and Clinical Psychology* requiring all submitted outcome studies to now include clinical significance analysis of some form (Crits-Christoph et al., 2013). Subject to extensions and adjustments (Christensen & Mendoza, 1986; Jacobson, Roberts, Berns, & McGlinchey, 1999), this entails an index reflecting a return to functioning within the non-clinical range (Jacobson & Truax, 1991; Jacobson et al., 1999; Kazdin, 1999). Clinical significance can be defined using a measure's clinical and nonclinical norms assuming that the norms reflect the individual or population being studied. Jacobson and colleagues (1999) provided three cut-off points to establish whether or not clients have demonstrated a shift from the dysfunctional to the functional range. Cut-off "A" is defined as the point two standard deviations from the mean of the dysfunctional population for the outcome measures used, in the direction of functionality. If a client's score exceeded this point, it suggested the client had moved from the dysfunctional range. Cut-off "B" reflects the level of functioning that fell within the range of the normal population. This range is recognized as starting two standard deviations below the mean of the normal population. The final cut-off point, cut-off "C", demonstrates the level of functioning that reflects a higher probability that the client was in the functional range as opposed to the dysfunctional range. When basing one's cut-off choice on sensitivity to change, Jacobson and colleagues (1999) indicated that when the dysfunctional and functional distributions overlap, cut-off A is the most stringent, cut-off B has the most relaxed criteria, and cut-off C, being the weighted mid-point between the dysfunctional and functional population, is deemed as the least arbitrary. Choosing the applicable cut-off is also based on the

availability of clinical and nonclinical norms, with Cut-off B being the cut-off used when only nonclinical data is available (Bauer, Lambert, & Nielsen, 2004).

Following this understanding of associated change with cut-off scores, the Reliable Change Index (RCI) is used to determine the magnitude of change, and whether or not this change is reliable (Jacobson et al., 1999). The RCI does this by establishing how large the difference needs to be between two scores taken from the same individual in order for it to be highly improbable that it is unreliable (Trauer, 2010a). This is calculated firstly by calculating the Standard Error of Measurement using the formula, $SE = SD\sqrt{(1 - r)}$, which combines the internal consistency (r) and the standard deviation of the measure used (SD). From this, the Standard Error of Difference is calculated (SE_{diff}) using the formula $SE_{diff} = \sqrt{2(SE)^2}$ (Jacobson & Truax, 1991). The RCI is then calculated by dividing the difference between the pre- (X_{post}) and post- (X_{pre}) treatment scores of an individual by the standard error of the differences between the two test scores (SE_{diff}) (Jacobson & Truax, 1991).

$$RCI = \frac{X_{post} - X_{pre}}{SE_{diff}}$$

If the resulting RCI is larger than 1.96, the probability of the mean difference occurring by chance is less than .05 (Jacobson & Truax, 1991).

The clinical significance method can describe meaningful outcomes that statistical significance may not, and ensure their applicability to the clinical setting. Clinical significance's promise lies within its versatility, relevant for studying both large groups of participants as well as in single-subject research designs (Hansen et al., 2002). Furthermore, not only does this method allow researchers to identify the overall rate of

change for clients who significantly improve, it can identify those who deteriorate significantly (Jacobson et al., 1999). Those who partially improve or yield no change are also identified. Jacobson and Truax (1991) extended previous methods, with clinical significance being applicable to measuring change within several diagnoses, whereas others focused only on a singular diagnosis (Jacobson & Truax, 1991). It also provides a much easier way to make comparisons between studies (Hansen et al., 2002). While this tool is useful, the interpretation of the index heavily relies on the use of appropriate measures for the client's impairment (Speer, 1992). Therefore, clinical significance can be achieved, but the findings are meaningless if the measures used do not reflect the symptoms for which the client has sought treatment (Speer, 1992). Clinical judgement is thus an essential tool to be used alongside the clinical significance method to ensure that the client not only demonstrates change as understood by outcome measures, but also in terms of meaningful change for the client (Fitzgerald, 2007).

Whilst preferable in clinical settings to statistical significance, clinical significance is not without its criticisms. Clinical significance can be seen as too stringent to be practical, as different variables can affect a client achieving meaningful change in the best of circumstances (Hansen et al., 2002). This also has an alternate criticism, with those who are less distressed at times failing to meet the criteria for being in the abnormal range in the beginning stages of therapy (Hansen et al., 2002). Despite this, clinical significance has become more frequently used in therapeutic outcome studies (Trauer, 2010c). Although alluded to, clinical significance should not be interpreted as evidence for meaningful change for the client, rather, that it produces a statistical measure that reflects reduction in symptoms (which may or may result in meaningful changes for a client). As such, clinical significant change does not necessarily equal meaningful change

from the client's perspective, but is currently identified as the closest statistical estimate available within psychotherapeutic research.

Where does the practicing clinician fit in?

Of course literal change could be assumed as something that a clinician can identify within session. However, research has indicated that clinicians appear to be overly optimistic in both identifying positive outcomes for their clients and their ability to achieve them. The positive optimism that clinicians appear to possess appears problematic due to the high incidence of negative outcomes that have been highlighted in research. Although positive outcomes are important, it should be expected that therapy does not work for everyone. Hansen and Lambert (2003) reviewed outcomes in 2,109 adults with varying disorders who were treated with one of 89 different therapies. Overall, these studies showed that two thirds of those who undertake therapy have a positive outcome by their 14th session, which then leaves a third of clients who either do not improve or worsen. Although there is an emphasis on the positive effect therapy has on individuals, research has also examined the rate of deterioration in clients. Estimates suggest that deterioration rates can be as high as 23%, although there are generally indicated as being between 5 and 10% (Kadera, Lambert, & Andrews, 1996). Up to 40% of clients have been estimated to not respond to therapy at all (Hansen et al., 2002; Lambert et al., 2002b; Newnham & Page, 2007; Parabiaghi, Barbato, D'avanzo, Erlicher, & Lora, 2005). While this finding is notable, it cannot be attributed solely to therapy alone, nor can it be assumed that therapy does no harm (Finch, Lambert, & Schaalje, 2001; Mohr, 1995). While clients may begin therapy with significantly deteriorated functioning, which can be linked to an intervention being less likely to have a positive impact, therapy may also contribute to negative outcomes by utilising longer than is

necessary strategies that are not effective for a particular client (Lambert & Shimokawa, 2011).

In an isolated study based on therapists' awareness of negative change, researchers used an archival client database of 13 licensed therapists whose clients deteriorated during routine care (Hatfield, McCullough, Frantz, & Krieger, 2010). Therapists had access to progress through therapy from the completion of the global symptom measure, the OQ-45, completed by clients prior to each session. Comparing clinician notes on indications of deterioration or improvement, 21% had some form of notes reflecting deterioration, with 59% having no mention of improvement. Only three percent were noted as improved. Out of the 41 who showed extreme deterioration, 32% were noted as deteriorating. This suggests that clinicians have some level of awareness regarding clients who do not improve, but identifying deterioration was not as developed even alongside the provision of outcome measures (Hatfield et al., 2010). Similar to the scientist-practitioner gap argument, this further suggests that these two methods, statistical and clinical judgement, both have limitations and thus can be used to supplement each other to provide clarity around client change (Dawes, 1996; Hatfield & Ogles, 2007; Lambert, 2010; Spielmans, Masters, & Lambert, 2006). Although the study indicated that routine outcome measurement did not prove to support clinicians in identifying deterioration, clinicians can be more aware of clients who are not improving as they should be by understanding trajectories of change. Recently, this has been thought to be of more assistance to practicing clinicians than understanding the overall outcome of therapy. Incidentally, this is the final feature of change.

The what and how of change

Evans' (2013) final feature of clinical change indicates that individuals progress through a pattern or sequence in achieving the goals of therapy. While a client may overcome their phobia of being in a social situation, how did the client get there? Did they first have to increase their exposure to the situation, or did they have to learn how to manage and reduce their anxiety first? Such an understanding follows the assertion that while understanding the effect of therapy is important, the understanding of change principles, that is, what clients respond to in therapy and how this impacts on positive progress (Carey et al., 2007), is more important clinically (Rosen & Davison, 2003). Understanding the shape of change gives insight into causal pathways of change, and being aware of the mechanisms that explain change in therapy is also likely to provide better structure to therapeutic systems and allow the quality of therapy to be maintained (Kazdin, 2009). Clinicians would then be better prepared to identify those individuals who aren't as susceptible to such strategies, and therefore more attentive to when changes need to be made to induce a positive response (Kazdin, 2009). The common factors considered essential to successful therapy can also be confirmed or identified, ensuring that these factors receive an appropriate amount of attention (Kazdin, 2009). The desire to know as much as possible about which therapy works best, for which group of problems, and knowing the number of therapy sessions that are required to achieve positive outcomes, is of added importance with psychotherapy changing to somewhat of a commodity, with those purchasing it wanting to know its effectiveness and value (Miller, Duncan, Sorrell, & Brown, 2005). Understanding this trajectory can also lead to improvements in understanding what needs to be assessed when identifying good outcomes in therapy.

In understanding how people change in CBT, however, is a symptom measure going to be enough? With a symptom being defined as “any deviation from normal functioning that is considered indicative of physical or mental disorder (American Psychological Association, 2007)” Evans (2013) proposes that emotional states, such as anxiety and depression, are both manifest and measurable in a number of different channels, for example, affective, cognitive, behavioural, and/or physiological components (Beck, 2011). At a basic level, one would assume mood and behaviour to be influential on change: again, using behaviour to confront difficulties, or reducing emotional state first? These different channels may not be alternative measures of the single construct that is “anxiety” or “depression,” but potentially independent expressions of the hypothetical constructs of each of these disorders. Each channel that makes up any given disorder construct is thought to have a reciprocal element, with improvement in one major problem area generally spreading into other areas. This can be understood in terms of “response relationships” first highlighted by Staats’ (1975), who indicated that behaviours/responses are often the independent variable for the emergence of other behaviours. Staats’ reasoned that many responses are arranged in a way that is cumulative and that hierarchical patterns signal that some behaviours/responses influence others (e.g. one needs to learn how to write before attempting poetry). Evans (2013) further elaborated that if there are a number of potential responses, all serving the same function within a repertoire, then fostering the desired alternative requires increasing the attractiveness of the currently less likely one. In terms of psychotherapy, the focus then depends on which aspect of a symptom results in a higher rate of improvement in the other aspects, thus providing insight for clinicians in treatment planning, particularly in terms of order of technique implementation, outcome evaluation, and facilitating enduring therapeutic change.

Based on the discussion around response relationships, to truly understand how clients arrive at a good outcome, solely focusing on overall symptom change is not likely to be enough. Taking into account the different channels of change, such as mood and behaviour, however, is likely to provide a greater understanding of how clients change. An attempt to understand different forms of change and how they progress throughout therapy was the subject of my Master's thesis (Fletcher, 2011). Asking both experienced and training clinicians to predict the course of change for both a depressed client and an anxious client undertaking a 12-session cognitive behavioural therapy protocol, change was predicted for each session via a mood, behaviour, and symptom measure. Guided by observations of graphs as opposed to statistical analyses, participants appeared to demonstrate simultaneous change between symptom and mood improvement. Behavioural improvement appeared to peak after symptoms and mood began to improve, with some clinicians predicting immediate behavioural change prior to symptom or mood change. While these findings have encouraged further interest in this area, this study was limited by the use of inconsistent methods in obtaining data and an absence of statistical analyses. As such, the current study aims to remedy such limitations and provide a clearer understanding of how these three variants of change progress throughout therapy.

Summary

In understanding the four features of change, that is, 1) identification of a construct to change, 2) evidence of change, 3) an endpoint which is based on a determinant of success and, 4) the pathway toward successful change, there are a myriad of ways to understand outcome. While symptomatic measures are predominantly used in identifying change, usually pre- and post- therapy, the shifting focus of research to understanding how clients change suggests that perhaps we need to also focus on the

varying channels of change, such as mood and behaviour. In doing so, we develop confidence in knowing that the change we are measuring is appropriate in terms of the goals of the client, as well as for our own clinical knowledge.

CHAPTER THREE: PERSPECTIVES ON CHANGE

Theories and Principles of Change

Now that a variety of ways to understand outcomes have been discussed, as well as acknowledging indications that the pathway to change is increasingly becoming the focus of research as opposed to overall treatment outcome, I will now move to how the process of therapy has been considered in research and how such an understanding can be implemented into clinical practice. A number of theories hypothesising likely pathways in which change is achieved have been established. Prior to discussing such theories, a definition of a theory of change is presented.

Defining ‘Theory of Change’

While change is of prominent importance in the psychological literature, it seems that the definition of what constitutes change is a topic of limited discussion (Gianakis & Carey, 2011). Furthermore, if the process of change is acknowledged, it is usually in terms of pre-post measurement (Carey, Carey, Mullan, Murray, & Spratt, 2006), and in terms of the clinician’s view, rather than a joint understanding from both client and clinician (Gianakis & Carey, 2011). From this perspective, it is not surprising that a definition of what constitutes a theory of change was not easily found in the psychological literature.

Theory has been defined by the American Psychological Association (2007) as “a principle or body of interrelated principles that purports to explain or predict a model of change” (p. 934). While also considering the term ‘model’ as a similar term, the definition of ‘theory’ can be extended to include “the representation of a concept...that

can be used for various investigative and demonstrative purposes, such as enhancing understanding of the concept, proposing hypotheses, [and] showing relationships..." (p. 586).

Therefore, for the purpose of this study, a "theory of change" is defined as a presentation of interrelated ideas and assumptions based on the expectations of how change is achieved, or what form of slope it represents in the psychotherapeutic process. Additionally, these theories should enhance the understanding of what change is, provide hypotheses based on how it occurs, and provide insight into the relationship between psychotherapy and change. This definition remains broad enough to identify theories that can later be narrowed down into more specific categories.

Linear change theories

When looking at what change is thought to "look" like, in terms of improvement plotted on a graph, linearity is the dominant idea in both research and clinical practice (Hayes, Laurenceau, Feldman, Strauss, & Cardaciotto, 2007b; Thompson, Thompson, & Gallagher-Thompson, 1995). It has been hypothesised that this predominant understanding reflects the methods used to measure change such as pre- and post-outcome and group mean statistics (Hayes et al., 2007b). In addition to linear changes being predominant, others have further detailed change as occurring in stages.

Some theories maintain that change can be understood as occurring in stages, with the first of such models being attributed to Carl Jung in 1944 (Jung, 1968) using psychotherapeutic and analytic methodology. Jung established four components essential to the "healing" process: confession, elucidation/interpretation, education, and finally transformation which Jung termed "analysis proper" (cited in Groesbeck, 1989). The first direct acknowledgement in the literature of a stage model was proposed by Carl Rogers

(Rogers, 1958), who established “stage sequencing”. This theory began with a process termed “loosening of feelings”, followed by the development of a new way in which one experienced “loosening of the cognitive maps experience”. This occurred alongside changing of the self and achieving congruence with experience. This model also involved shifting from ineffective to effective choice, freeing oneself from relationship fears, and developing sharp differentiation between feelings and meanings (Rogers, 1958).

In terms of stage models that address symptoms directly, a two-stage model was developed by Unlethuth and Duncan (1968) which sought an explanation for clients whose symptoms dramatically improved in the first four weeks of treatment (Uhlenhuth & Duncan, 1968). This stage model posits that initial change originates from placebo or nonspecific effects, but their rate of change is influenced by the client’s expectations and responsiveness. The second part of this model proceeds in a gradual manner, influenced by the understanding of the client’s symptoms and how they have been established. In more recent years, the Stages of Change and Three Phase models are two stage theories that have received attention in the psychological literature.

The Stages of Change model (Prochaska & DiClemente, 1982) has been influential in psychotherapeutic discussions, especially in terms of motivation. This theory proposes that clients present to therapy in one of five stages of motivation: pre-contemplation, contemplation, preparation, action, or maintenance. *Pre-contemplation* pertains to an individual who has little motivation or commitment, and has no intention to seek any change to their current situation in that moment, or in the near future. The stage of *contemplation* follows, with an individual acknowledging that a problem exists with intention to address it within the next six months. This is followed by the *preparation* stage which reflects an individual who has sought assistance in making changes to issues

they are currently experiencing. The *action* stage reflects an individual who has become involved and committed to addressing the problem, focusing a substantial amount of time and energy to achieve change. Lastly, the stage of *maintenance* entails the resolution of the issues, with an invested effort to maintain change that has occurred to avoid relapse. The problem is considered to be maintained following six months of reduction in the client's impairments. Clients may vary in the length of time they spend in each stage, but the strategies remain the same (Norcross, Krebs, & Prochaska, 2011).

To evaluate the effects of the SOC model in psychotherapy, a meta-analysis looked at studies based on face-to-face psychological therapy provided by mental health professionals (Norcross et al., 2011). The studies included clients who had a DSM-III or IV diagnosis, and incorporated therapy lasting more than three sessions. Out of the 1,686 applicable studies, only 39 (2%) met the criteria for a medium effect, which demonstrated the SOC's ability to reliably reflect the sequence of progression through motivational phases across the course of therapy. This small percentage of supportive studies suggests that while the SOC model is highly regarded, the actual progression through the stages is complex to analyse in terms of its impact. For clients who present to therapy committed to making changes and have high motivation and high awareness of their problems, therapy is likely to be brief (Steenbarger, 1994).

While this model touches on the active stage of therapy and subsequent resolution of difficulties (action and maintenance stages), the stages pertaining to the commitment of the client to therapy appear to receive the most attention. As such, despite the name, this theory appears to have less to do with how client's change in therapy, focussing more on how a client commits to change (Evans, 2013). With the limited discussion and support

base of why a linear pattern of change is expected, it is no surprise that nonlinear patterns are emerging as the likely pattern of psychotherapeutic change.

Nonlinear and discontinuous patterns of change

Only recently has the idea that therapeutic change does not progress in a linear fashion been entertained by researchers (Kazdin, 2007), despite being acknowledged in earlier studies carried out by Howard and colleagues' (Howard, Kopta, Krause, & Orlinsky, 1986), Barkham and colleagues (1993) and Tang and Derubeis (1999b). Nonlinear and discontinuous patterns have been further developed and promoted by the growing emphasis on the use of routinely monitoring change across therapy (Stiles et al., 2003), and an increased need to understand individual trajectories of change (Thompson et al., 1995). Furthermore, studies have provided preliminary evidence that change in therapy often occurs during critical incidents such as when anxiety has been addressed (Grafanaki & McLeod, 1999; Paulson, Truscott, & Stuart, 1999; Timulak & Elliott, 2003). Represented by large changes in affect, these are meaningful points in therapy where a client comes to a realisation specific to their problems, followed by a collaborative effort to work through the problem, and ending with resolution (Greenberg, 2007).

While not having a primary interest in studying a smooth pathway of change with little variation, there are a number of sources where non-linear patterns have been demonstrated. With pre-post measures as the predominant way of assessing change, multiple trajectories for the achievement of client change may be obscured (Barkham et al., 1993). Similarly, research typically focuses on group averages as opposed to individual statistics which may also provide information that is not representative of the individuals being studied (Hayes et al., 2007b). Despite these potential problems,

Steenbarger (1994) found, while reviewing the dose-effect relationship, that change tends to occur in waves, with symptoms increasing and decreasing throughout the therapeutic process. Another study looking at changes in depressed clients showed that those who best responded to therapy overall were those who experienced a “bumpy ride” (Thompson et al., 1995). Likewise, Kadera and colleagues (1996) showed that patients who positively improved rarely did this in a steady, linear fashion.

It is not surprising then, that a number of recent theories developed explain nonlinear and discontinuous patterns that arise across the psychotherapeutic process, with the three phase model, the assimilation model, developmental change theory, and dynamical systems theory being of particular interest.

Three phase model. In their renowned study, Howard and colleagues (Howard et al., 1986) evaluated the relationship between therapy duration (dose) and outcome (effect). Interest in this area was first inspired by Seeman (1954), with earlier studies suggesting a positive association, revealing a weak curvilinear relationship consisting of a negative decelerating slope (Cartwright, 1955; Johnson, 1965; Miller & Berman, 1983; Smith, Glass, & Miller, 1980; Strassberg, Anchor, Cunningham, & Elkins, 1977; Weitz et al., 1975). This decelerating slope has generated a moderate amount of support (Finch et al., 2001; Hansen & Lambert, 2003), and has been hypothesised as being due to initial sessions having the most effect, with the remaining sessions decreasing in effect as therapy progresses (Barkham et al., 2006). An alternative reason for this shape of change reflects the idea that those in therapy who reach their own good enough level of improvement opt out of continuing therapy (Barkham et al., 2006).

Following the initial study, Howard and colleagues (Howard, Lueger, Maling, & Martinovich, 1993) shifted their focus to developing a stage model designed to represent

the typical process of client change in psychotherapy. The resulting effort was the Three Phase Model, depicting the stages of remoralization, remediation, and rehabilitation. The stages are sequential, with improvement required to occur in the earlier stage before moving on to the next.

Remoralization. Individuals presenting to therapy in the remoralization stage are characterized by hopelessness and desperation (Howard et al., 1993). Their issues have become all consuming, leading to difficulties in enacting their typical coping strategies, resulting in severe impairment to the client's functioning. The alleviation of the client's subjective wellbeing is thus the main focus of the remoralization stage. This is the shortest phase in this model, with movement to the next stage typically occurring in a few sessions.

Remediation. This stage focuses on reducing the issues the client has sought therapy for (Howard et al., 1993). This can be in relation to symptomatology, life problems, or both. Both client and therapist address these issues by collaboratively developing and enacting coping skills, with an emphasis on building skills that operate more effectively than those that the client is currently using or has previously used. This stage can be very structured, and is suggested to take approximately 16 sessions to complete.

Rehabilitation. Clients use insight into the origin of their problems to determine what pervasive patterns and factors are linked to the development of their issues, better preparing them should these patterns arise in the future. This stage can take a lengthy amount of time, dependent on the severity and chronicity of the presenting problem.

Analysing self-reports from a large sample of clients at intake (463), session 2 (182), session 4 (157), and session 17 (73), Howard and colleagues (1993) produced findings showing that clients undertaking psychotherapy progressed in a negative, decelerating slope in line with remoralization, remediation, and rehabilitation. A number of studies have also supported the consistency of the phase model in clients, as well as the decelerating curve proposed by the dose-effect therapy which this model promotes (Barkham et al., 1996a; Callahan, Swift, & Hynan, 2006; Lutz et al., 2001).

Assimilation model. Heavily guided by behavioural theory, the assimilation model was originally developed to create a linear model reflective of statistical predictions of psychotherapeutic change. However, the authors instead identified non-linear patterns occurring through the use of behavioural techniques, such as exposure therapy (Stiles et al., 2006). Presented as a stage model, clients progress through different stages regarding their problematic experiences (Stiles et al., 1990). Each stage represents the extent of integrating problematic experiences within the client's own framework. While having a behavioural influence, this model is thought to be applicable across multiple theoretical orientations. It is also considered to be similar to the Stages of Change (SOC) model for three reasons; it is not essential for clients to begin therapy in the first stage; the process is sequential; and clients can revert back to a stage they have already completed (Barkham, Stiles, Hardy, & Field, 1996b).

The purpose of therapy is to assimilate and accommodate a client's problematic experiences with more helpful, appropriate ways of experiencing a particular situation. Identifying a client's stage has been made easier with the development of the Assimilation of Problematic Experiences Scale (APES), describing the process in a sequence of eight stages (Stiles, 2001). Represented as a client's internal voice, the

assimilation model begins with the stage *warded off* where clients are completely unaware of the issues that are causing their impairment. This is characterised by minimal expression or successful avoidance. The *unwanted thoughts* stage follows, characterising a client who has not acknowledged their issues but exhibits psychological discomfort. Expressed as acute psychological pain, *vague awareness/emergence* describes an individual who cannot connect their problems and discomfort to a specific experience. The fourth stage concerns *problem statement/clarification* where the client and therapist begin identifying potential origins of the problem, and proceed to introduce the resolution of these issues as goals. Individuals then proceed into the *understanding/insight* stage, where a client's affect becomes mixed as a response, a collaborative effort in determining the schema reflecting the client's issues, and the formulation of connective links. *Application/working through* follows with the understanding developed in the previous stage used as a basis for improvement, along with an emphasis on optimism and hope. The next stage, *problem solution*, entails successful achievement of goals alleviating psychological pain and promoting positive affect. The last stage reflects *mastery*, where generalisation of this newfound knowledge and understanding is explored, enabling the client to enact such strategies in future situations.

The assimilation model is characterised by fluctuations across therapy reflected by variability in the intensity of a client's distress and symptoms (Stiles et al., 2006). This is acknowledged by the assumption that those who enter therapy within the *warded off* or *unwanted thoughts* stages typically worsen before improvement based on potential resistance such clients have in confronting problematic experiences that have led them to seeking therapy (Stiles, 2006). Additionally, negative mood is suggested to peak during the *vague/awareness* stage (Stiles, 2006). Clients are encouraged to acknowledge and understand their problems from *problem statement/clarification* to *mastery*. As a result,

positive affect improves over these stages, with the most rapid improvement expected between *application/working through* and *problem solution*. Clients are expected to conclude the *mastery* level in a functional range reflective of the normal population (Stiles, 2006).

In the empirical validation of this model, case studies have been the primary method of study. This qualitative approach examines a single client, or a series of cases, progressing through this model, rather than completing a randomised clinical trial with numerous participants (Stiles et al., 1991). Despite this, a range of symptom patterns have been evaluated including depression (Brinegar, Salvi, & Stiles, 2008; Caro Gabalda, 2006; Field, Barkham, Shapiro, & Stiles, 1994; Honos-Webb, Stiles, Greenberg, & Goldman, 1998), anxiety (Stiles, Meshot, Anderson, & Sloan, 1992; Stiles et al., 1991), and personality disorders (Humphreys, Rubin, Knudson, & Stiles, 2005; Osatuke et al., 2005). Interestingly, when comparing those who responded positively to therapy with those who did not, following a process consistent with the assimilation model resulted in positive outcome.

Developmental change theory. Originating from developmental theorists, the developmental change theory applies aspects of the interactions between child and parent to the psychotherapeutic process in explaining a client's experience (Stern et al., 1998). Consisting of four stages, developmental change theory begins with the applied evaluation of Tronick and colleagues' (1998) description of the parent-infant interactive process. Involving the matching and mismatching of ruptures and repairs in the relationship, *moving along* involves the patient and therapist working together towards a common goal (Stern et al., 1998). Both clinician and client understand that the goal can change, and the pathway to achieving the goal is virtually unknown. There can be a

variable number of steps in the *moving along* process, with each step being known as a *present moment*. These moments are short periods of interaction that occur within that particular moment, between the client and therapist, representing redefining and redirection of the therapeutic topic towards the goal in question (Stern et al., 1998). Acknowledging that while changes in the definition of the goal can occur, improvement in this stage is hypothesised to be largely linear (Stern et al., 1998).

The second stage concerns *now moments* (Stern et al., 1998). This reflects an unpredicted emotionally charged moment of truth, which produces a completely different viewpoint to what the client had believed prior to that precise moment. Both client and therapist are drawn fully into the present, which threatens the stability of the client's state prior to this realisation. Furthermore, this "aha" moment can lead to a transition into a new state of reorganization. Due to its unpredictability, and the intensity of such a moment, this can represent a non-linear jump in the client's experience (Stern et al., 1998).

While some *now moments* may be missed, those that are recognized and seized therapeutically can become a *moment of meeting* (Stern et al., 1998). Based on the unexpectedness of the *now moment*, a *moment of meeting* entails participants making an on the spot, individual response that is both unique and authentic. Notably of importance in this *moment of meeting* is that it must arise from the clinician's own sensibility and experience, and not the result of habit, technique, and theory (Stern et al., 1998). For both the *now moment* and the *moment of meeting*, the parent-infant interaction has been shown to exhibit comparable movements in behaviour and inter-subjective states (Sander, 1994).

Based on the infant-parent observations of Sander (1988), an *open space* has been suggested to occur immediately after a *moment of meeting*. It entails those involved in the

relationship being able to disengage during a meeting, where they can be alone yet still be in the presence of the other. In psychotherapy, this is thought to be the point where a client can integrate the *moment of meeting's* effects and become at peace with the new inter-subjective state. This is followed by the reuptake of *moving along* now done through the new state (Stern et al., 1998). As such, change is characterised by fluctuations in symptoms and offers some suggestion that individuals become worse before getting better.

Dynamical systems theory. Similar to the model proposed above, dynamical systems theory views change as a reorganisation of a system (Hayes et al., 2007b). Put simply, a dynamical system is the interaction and continuous process of a set of factors over time (Vallacher, Read, & Nowak, 2002). Thus, dynamical systems theory is the application of mathematical theory reflecting this process (Hayes et al., 2007b).

Dynamical systems theory was developed based on “perturbation experiments”, a method which involves causing disturbance within a system, followed by the frequent measurement across the given time period of variables of interest (Hayes et al., 2007b). With a baseline measured, a disturbance is prompted and the process of change is observed, identifying facilitators and barriers. Perturbation studies examining psychotherapy identified “critical fluctuations”, an important discontinuity suggestive of transition (Kelso, 1995). Critical fluctuations imply that when introducing new elements change is not gradual and linear as information becomes too great for a system to make sense of (Kelso, 1995). Rather, the system suffers severe disturbance, resulting in instability before the system can reorganise itself. Alongside this system destabilization, the system receives an increasing amount of new information, resulting in the search for better fitting patterns to enable the system to function better. Old, less practical patterns

then compete with new patterns, resulting in a new dynamic state based on the combining of old and new patterns. These critical fluctuations have been found to precede affective, cognitive, and interpersonal changes within individuals in psychotherapy (Vallacher et al., 2002), and have resulted in major meaningful transformations in one's life (Baumeister, 1991; Linley & Joseph, 2004).

When referring to psychotherapy, therapists typically disturb the client's current way of functioning and facilitate important life changes by interrupting, challenging and destabilizing old, unhelpful patterns (Hayes et al., 2007b). Dynamical systems theory thereby presents a way of understanding how techniques used in psychotherapy create change, and what kind of shape this change would take. While suggestive of non-linear patterns, this model also acts under the assumption that change can be linear for certain clients and does not assume that there is one all-encompassing change pattern (Hayes et al., 2007b).

Statistical predictions and feedback

As theories of change become more developed, how they can be used becomes clearer. Stemming from the Patient-focused research methodology, and driven by the rate of non-responders of therapy, Lambert, Miller, Lutz, Howard and other psychotherapy researchers set about developing a statistical prediction method which could assist practicing clinicians in identifying clients who were not responding. This was based on the premise that client outcome is highly predictable using statistical methods (Brown, Dreis, & Nace, 1999; Haas, Hill, Lambert, & Morrell, 2002; Hansen & Lambert, 2003; Howard et al., 1996; Miller, Duncan, & Hubble, 2004; Speer, 1992; Steenbarger, 1992).

Wanting to implement session-by-session routine outcome measurement, researchers focused on developing an outcome measure that was attuned to areas thought

to be of importance when understanding psychotherapeutic change (Lutz et al., 1999). This process began with the creation of the Outcome Questionnaire-45 (OQ-45), a measure designed to be administered no more than once a week, prior to each session, that accounted for symptomatic functioning, interpersonal problems, and social role performance (Lambert, 2010; Lambert et al., 2001b). A shorter version, the Outcome Questionnaire-30, has also been created. In addition to these measures, COMPASS, a computer programme that assists in predictions and reports using scores from the OQ-45 (Lueger et al., 2001), has also been developed.

Following the development of an appropriate measure to assess client improvement over time, the first attempt to predict client change statistically was established. Termed the “rationally derived method”, two clinical judges used their knowledge about early response to treatment, the dose-response effect, and clinically significant change to devise algorithms that were able to predict negative outcome (Lambert, 2010). Using the OQ-45, the client’s initial severity, amount of change that had occurred at the session of measurement, and the number of sessions of care an individual had received, a matrix was generated that identified individuals who were not demonstrating progress. Plotting the severity (i.e. initial OQ-45 score) on the vertical axis, and OQ-45 scores expected for sessions of interest on the horizontal axis, judgements were made on what score would warrant a satisfactory response, a return to normal functioning, questionable progress, or a highly problematic response. Green, white, yellow, and red dots represented these responses respectively and were devised based on the client’s individual initial score. A white indicator suggested a client fell within the functioning range (below a score of 64 on the OQ-45), and that termination should be considered. A green indicator was suggestive of the client progressing adequately, the yellow indicator suggestive of a client who was responding in a less than adequate

manner, and red being indicative of a client who was not making the expected levels of progress (Lambert, 2010). To minimise the complexity of allocating a predicted OQ-45 score to each session but still present therapists with adequate information to be alerted to negative progress, the therapeutic process was divided into three timeframes; session two to four, five to nine, and ten and above.

Following the creation of these matrices and the repeated administration of the OQ-45, the accumulated data enabled the development of an empirical method of prediction. Termed the “expected treatment response model” (ETR, Lambert et al., 2003; Lueger et al., 2001), researchers applied hierarchical linear modelling (HLM), involving the statistical analysis of factors that vary at different levels, to client data to define the shape and rate of change a client is expected to follow session-by-session over the course of therapy based on their initial level of severity (Finch et al., 2001; Lambert, 2010). HLM was used to identify the shape/rate of change using a two-level model, firstly addressing a Mental Health Index score which reflected the intercept, slope, and random error for each client’s change trajectory. The second level concerned identifying predictors of the level 1 intercepts and slopes (Lueger, Lutz, & Howard, 2000). Using Jacobson and Truax’s (1991) indexes of clinical significance, clients were categorised by the severity level of their mental health functioning.

The main intention for ETR in clinical practice is to inform and supplement clinical decision making. While this initial method has been successful in achieving that, it is the hope of the designers that each clinic aggregates their own data to generate information specific to their future clinical population (Lueger et al., 2000; Lutz et al., 2006b). ETR was designed to help clinicians develop and improve by applying findings from studies to their accumulated clinic data (Lutz et al., 2006b), reflective of the desire

of patient-focused research to minimise the scientist-practitioner gap. Subsequent to an understanding of improvement, it was also a primary interest of researchers to generate recovery curves and produce tolerance intervals that alerted clinicians to the approximately 10% of clients who deviate from the pathway of positive outcome found within this research (Finch et al., 2001). While New Zealand is one of the front runners in the use of routine outcome measurement, it appears that this is in the context of service management, as opposed to for the use of understanding how clients are likely to progress in therapy (Trauer, 2010c).

Following the achievement of empirical support for the ETR method for the OQ-45 and associated variations (Bishop et al., 2005; Bybee, Lambert, & Eggett, 2007; Ellsworth, Lambert, & Johnson, 2006; Hannan et al., 2005; Lambert et al., 2002a; Lutz et al., 2006a; Lutz et al., 1999; Spielmans et al., 2006), ways in which therapy could be enhanced by predictive models were identified. Primarily, these predictive models were found to be useful alongside the provision of feedback to clinicians; information on whether the client is progressing positively, showing no change, or deteriorating. When first attempting to address the effect of feedback provision, the main benefit found was maintenance of client attendance in those who were not responding to therapy as expected (Lambert et al., 2001b). Clinicians who received feedback typically had clients who achieved better outcomes and completed approximately double the sessions (Lambert et al., 2001b) with these findings having been replicated (Lambert et al., 2002b). While feedback has been shown to improve the functioning of those who are deteriorating in therapy, evidence has demonstrated some findings that feedback has little effect on non-responders (Hawkins, Lambert, Vermeersch, Slade, & Tuttle, 2004).

The use of ETR and feedback in routine practice provide a number of opportunities. A formal evaluation of the expected effectiveness any therapeutic approach may have on an individual would be provided along with information regarding variability between clients, allowing for the study of certain characteristics that impact on the rate at which a client benefits from therapy (Newnham & Page, 2010). With outcome measures administered at each session, clients who deteriorate can be accounted for in a quick manner, with more information available in comparison to the use of pre- and post-therapy measures (Lutz et al., 1999). ETR and feedback better prepares clinicians for clients who are most at risk of treatment failure (Newnham & Page, 2010). It can provide clinicians with valuable information regarding their practice and therefore upholds the ethical responsibilities imposed on clinicians and service providers, promoting accountability (Newnham & Page, 2010). In terms of research specifically, the grouping of common characteristics can be useful in identifying specific responses to treatment (Lutz et al., 1999). The ability to identify and make empirically-based decisions in real time for those who are responding to treatment negatively is another advantage. Additionally, feedback provision is also directly helpful for the client. Not only does it provide clinicians with another level of reference in evaluating treatment response, it enables the client to visibly see their progress, encouraging them to be more involved in their own care (Newnham, Hooke, & Page, 2010).

The literature on the benefits of feedback is a work in progress, with findings so far tentatively showing that the provision of feedback reduces negative outcomes in psychotherapy for a number of problem areas including mood and anxiety disorders (Harmon et al., 2007; Hawkins et al., 2004; Lambert et al., 2001b; Lambert et al., 2002b; Whipple et al., 2003). While the research basis is not entirely clear on the value of feedback in terms of CBT and routine clinical practice specifically, it can provide the

long-term benefit of growth and development for clinicians, and the short-term result of being informed about cases requiring additional attention (Sapyta, Riemer, & Bickman, 2005).

Summary

With the continual improvement and clarification of research methodologies, the more sophisticated and attuned to client change clinicians can become. As can be seen from the theories discussing change so far, change has been understood as a linear trend, a decelerating curvilinear trend, as well as having some non-linear aspects to the pathway towards change. As the literature on such theories becomes more substantial, how this information can be used will be better understood, including understanding how the accumulation of such data can suggest how a treatment impacts on a client's progression towards change – thus suggesting when and where particular aspects of client change need be targeted over the course of therapy. Furthermore, this data can then be applied to clinical practice to help identify those clients who are not responding to an empirically based treatment as expected.

Other Aspects Related to the Pattern of Change

It is clear from the literature that the focus on understanding change is important in identifying the true value of psychotherapy treatments, along with understanding what kind of change would be expected of an individual and at what point in therapy. Having such information can help us determine what areas of change we need to monitor or measure to understand whether change is occurring by telling us at what point in therapy it occurs. Furthermore, it can alert us to clients who are not progressing in a manner that

is likely to result in a positive outcome. So, from the research that has been completed so far, what does change look like over the course of therapy?

An approximate linear decline in symptoms is often suggested in the literature regarding change (Barkham et al., 1996a; Thompson et al., 1995). However, as has already been identified, it has been suggested numerous times that therapy produces a negative progressive slope or a negative decelerating curve for adverse symptoms as they decline (Howard et al., 1986; Stulz, Lutz, Leach, Lucock, & Barkham, 2007; Thompson et al., 1995). When reflecting on the differences between mood and anxiety disorders, Howard and colleagues' (1986) original article showed that depressed patients respond quicker to therapy than those who have an anxiety disorder, though still representing a decelerating curvilinear trend (Howard et al., 1986). As non-linear patterns reflecting dynamical systems theory are further explored, particular patterns that may occur in psychotherapy may be identified. Three of these are early rapid response, the sudden gain, and the depression spike.

Early rapid response

In studying the early rapid response, Ilardi and Craighead (1994) found that 60-70% of client improvement occurred within the first four weeks of therapy, followed by change progressively levelling off. In response to controversy, a number of theories have been proposed to explain an early rapid response. One idea is that clients demonstrate individual differences in rates of responding when compared to those who respond early and those who do not. This has been suggested as being related to a client's preparedness to change (Lambert, 2005). Other theories propose that, instead of being a response to specific treatment factors, clients who respond early to therapy can be seen across a number of theoretical orientations, suggesting that common factors across numerous

therapies may be responsible (Ilardi & Craighead, 1994; Lambert, 2005). This idea stems from the defined period this type of improvement typically occurs in prior to the implementation of specific therapeutic techniques suggesting that it is too soon for effects to begin taking place. Agreement on why early rapid responses occur has not been achieved, with some suggesting that it is an alternate form of other discontinuous patterns, such as the sudden gain, where a client demonstrates a significant improvement between one session and another (Tang & DeRubeis, 1999). Studies regarding rapid early responses typically focus more on how these patterns impact on outcomes as opposed to identifying the number of clients who typically experience them. Of the few studies that do offer this analysis, early rapid response patterns have been shown to occur in 25 to 41% of clients (Hayes et al., 2007b; Johnson, Lambert, & Sullivan, 2001).

Understanding why this phenomenon occurs has been made further problematic by the lack of consensus on what defines an early rapid response. Lambert (2005) presented a list of various attempts to define early responding including the use of a median split (Stewart et al., 1998), using clinician ratings of change between two weeks (Present et al., 2008), whether or not clients exhibit little or no psychopathology following the first session of therapy (Stewart et al., 1998), as well as those clients who demonstrate 50% of their overall change over a pre-specified number of sessions (Beck & Steer, 1990; Meites, Ingram, & Siegle, 2012). Issues arise with the use of weeks as the defined period in which early rapid response occurs, rather than number of sessions (Lambert, 2005). Tang and DeRubeis (1999a) suggest that a set time period is problematic as number of sessions and weeks represent different scales of measurement. Dependent on the client's needs, sessions can be set up to occur bi-weekly, weekly, or fortnightly, although different variations can also take place. Subsequently, this position suggests that some specific techniques may be observable if the defined time period of an

early rapid response is based on number of sessions rather than weeks (Tang & DeRubeis, 1999a).

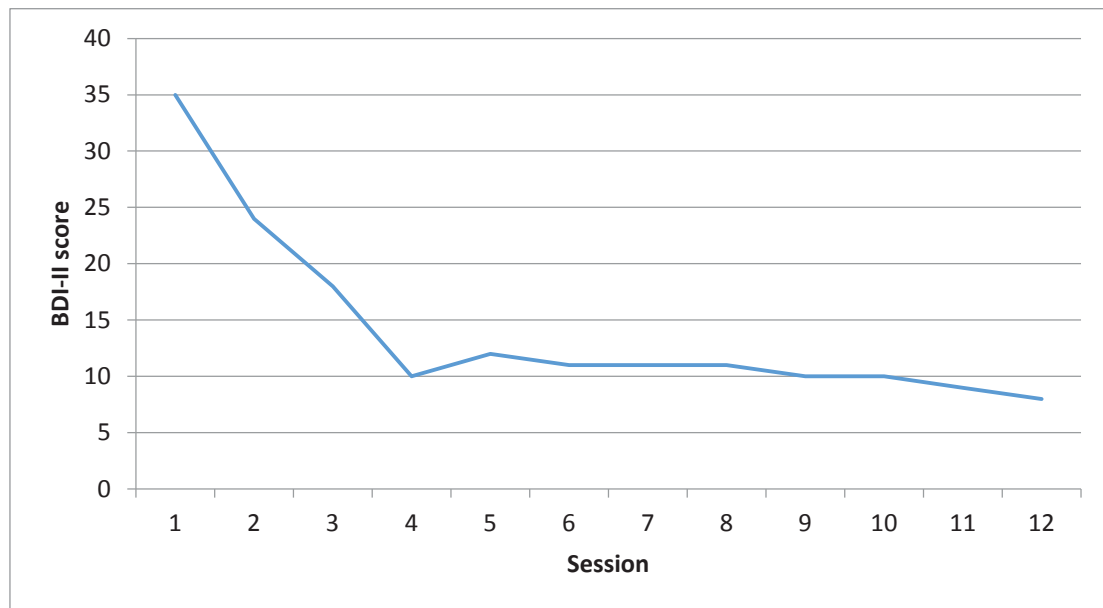


Figure 3.1. Example of an early rapid response using Beck Depression Inventory scores.

The absence of a comprehensive definition complicates the discussion on what early rapid response actually represents (Lambert, 2005). Haas and colleagues (Haas et al., 2002) combined session-by-session scores in relation to normative expected responses to develop an applicable definition. For identifying those who demonstrate an early response, this equation was proposed: the mean difference between the overall “expected recovery” and the “actual” recovery across sessions 1 to 3, divided by the average representing more than 65% of the expected recovery, with the “expected” recovery derived from the data of 10,000 clients who had undergone repeated measures of the OQ-45 (Haas et al., 2002). While based on a singular measure, such concepts have been shown to be easily adjusted to other measures, although accumulation of data is required (Hayes et al., 2007b). Therefore, as such an equation requires data of numerous clients, it is difficult to apply in terms of general research and the application of Ilardi and Craighead’s 60% -70% of change occurring is often used.

The long-term benefits of an early rapid response occurring in therapy have been identified in a literature review which concluded that those who respond early in therapy are likely to have enhanced positive end of treatment and long term outcomes (Lambert, 2005). Research focusing on types of early response has found three typical patterns (Lutz, Stulz, & Köck, 2009). One study found that clients presenting with moderate to severe depression demonstrated either moderate early improvement (mean BDI score decreased by 9.7 within the first 8 weeks of therapy; early response effect size of 1.33) or rapid early improvement (mean BDI score decreased by 18.1 points; early response effect size of 2.90), and those with mild to moderate depression exhibited moderate early improvements (mean BDI score decreased by 5.3; early response effect size of 1.05). These improvement patterns were predictive of outcome at 18 month follow up (Lutz et al., 2009), with early rapid responding again predicting a continued positive response to therapy. In a study aimed at understanding discontinuities in CBT treatment of depression and anxiety, findings showed that those who responded early to therapy demonstrated more hope in the initial sessions (Hayes et al., 2007a). The negative decelerating curve of the dose-effect model further offers support to the early rapid response literature with clients showing greater response to therapy in earlier sessions (Kopta, 2003).

Sudden gains

In the defining sudden gains study (Tang & DeRubeis, 1999b), the authors hypothesised that clients can experience large decreases in depressive symptoms at a rapid rate, and that a decrease in depressive cognitions was responsible for such a response. According to Tang and DeRubeis (1999b), to define a sudden gain three quantitative criteria must be met. Firstly, it is essential that the sudden gain is large in absolute terms, understood through clinical significance. Specifically, for depression, this

was defined as at least a decrease of seven Beck Depression Inventory points based on empirical justifications completed in prior research identifying a spike of 7 points or more as being qualitatively different from smaller ones (Tang, DeRubeis, Beberman, & Pham, 2005). Secondly, the gain must represent at least 25% reduction of impairment in the session prior to the gain as identified by scores on a reliable and valid outcome measure. Thirdly, the three sessions preceding the gain must have a mean level of functioning significantly lower than the three sessions following the gain (Hofmann, Schulz, Meuret, Moscovitch, & Suvak, 2006). Furthermore, in one between-session interval, a sudden gain can be representative of 50% of the total improvement a client achieves in therapy (Tang et al., 2005). Sudden gains have also been shown to reverse. This occurs when clients lose 50% of the improvement that was attained during the sudden gain (Tang et al., 2005).

In studying the generality of sudden gains in 135 clients, 17% of clients experienced an early sudden gain as opposed to the original 39% established in Tang and DeRubeis' 1999 study (Stiles et al., 2003). Upon examination, findings seem to suggest that sudden gains may be one of many fluctuations seen in therapy, although the most intense and sustained fluctuation (Stiles et al., 2003). Studies on sudden gains have mainly focused on depression, with one study based on sudden gains in social phobia showing that while clients do not meet the strict criteria of the sudden gain, they typically experience fluctuations in their experience of symptoms (Hofmann et al., 2006).

As with early rapid responses, sudden gains have been shown to be predictive of positive therapy outcomes, with findings demonstrating that those who experienced sudden gains showed significantly less depressive symptoms when treatment finished and at 18 month follow-up (e.g. Gaynor et al., 2003; Hardy et al., 2005; Stiles et al., 2003;

Tang & DeRubeis, 1999b; Tang et al., 2005; Tang, Luborsky, & Andrusyna, 2002; Vittengl, Clark, & Jarrett, 2005). While researchers suggest that findings are attributable to reduced negative cognitions, sudden gains are not explicitly attributed to CBT. They have been found in brief dynamic therapy (Tang et al., 2002), non-manualised psychotherapy (Stiles et al., 2003), and systemic behavioural therapy (Gaynor et al., 2003), which found sudden gains in between 17 – 39% of clients. Furthermore, in a study aiming to examine the validity of sudden gains, findings showed that they occurred in both placebo pill and pharmacotherapy treatment alongside clinical management (Vittengl et al., 2005). This finding suggests that sudden gains may not be a completely unique result of psychotherapy and, like early rapid responses, may be due to readiness of clients and other common factors (Vittengl et al., 2005).

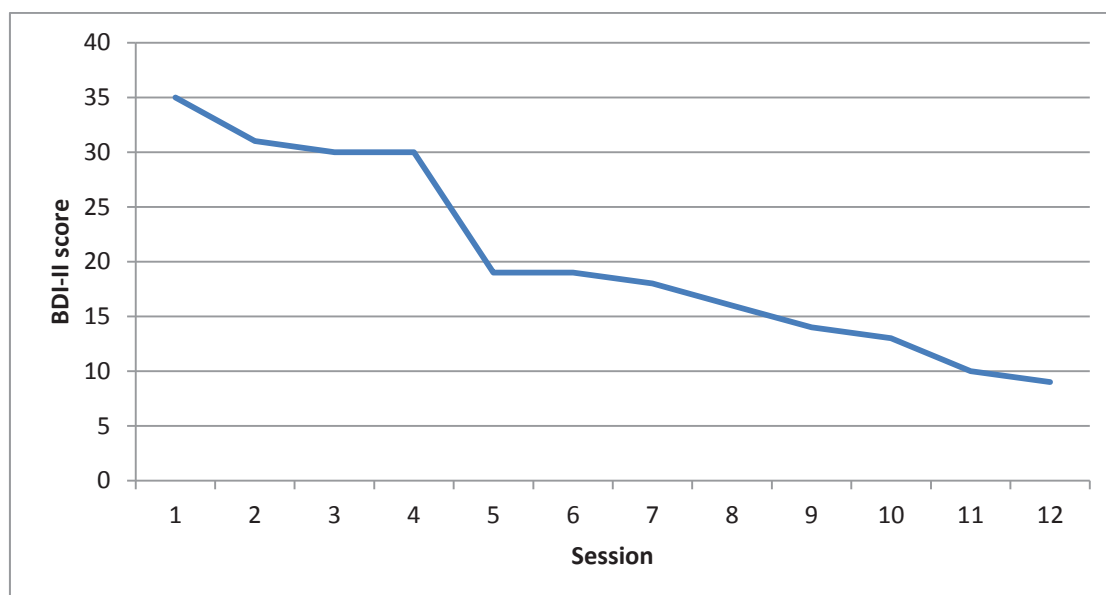


Figure 3.2. Example of a sudden gain using Beck Depression Inventory scores.

Depression spike

One of the more recent non-linear concepts presented in the empirical literature is the depression spike (Hayes et al., 2007b). This concept refers to a considerable increase

in depressive symptoms, followed by a considerable decrease. To define a depression spike, researchers have used similar criteria to sudden gains, although in reverse terms (Hayes et al., 2007a). That is, a depression spike follows a significant increase in depressive symptoms, indicated by a seven point increase or more on the BDI, followed by an equivalent decrease (Hayes et al., 2007a). This then depicts a change progression over therapy described by a cubic function, that is, an increase in symptoms followed by a further decrease. Based on this criteria, Hayes and colleagues (Hayes et al., 2007a) found that depression spikes occurred in 62% (18) clients in their study. Notably, additional studies have replicated such spikes using similar measures of the same construct suggesting generalisability of this phenomenon (Hayes et al., 2007a). While a relatively new line of research, studies have suggested that the occurrence of depression spikes do not specifically occur in early sessions, but typically take place prior to session 11 (Hayes et al., 2007a). When looking at the depression spike in terms of emotional-cognitive processing, one study found that those who experienced depression spikes experienced higher processing during the period of arousal than those who did not demonstrate a spike (Hayes et al., 2007a). The depression spike also showed a heightened probability of positive outcome.

While these three non-linear patterns, early rapid response, sudden gains, and the depression spike, have typically been researched in clients with depression, research has also identified similar discontinuities with anxious clients (Hayes et al., 2007b; Heimberg & Becker, 2002; Tang & DeRubeis, 1999b). The majority of these studies have focused on emotional arousal, which is induced by exposure-based techniques (Hayes et al., 2007b). Disturbance is caused to the client's "system" by provocation of fear stimuli. As explained in dynamical systems theory, new ways of responding to the stimuli are then explored and reorganised, creating a more adaptive response. When using exposure

techniques, patterns of change have been identified (Heimberg & Becker, 2002). Similar to the suggestion proposed by numerous studies, a steady gradual linear decline has been found in clients being treated for anxiety. A spike similar to the depression spike has also been found, occurring directly after the initial exposure to the fear construct. Termed the “habituation curve”, this non-linear pattern proposes a brief reduction in symptoms, similar to the early rapid response or sudden gain (Heimberg & Becker, 2002). It is then followed by a sharp increase in symptoms to the highest level of distress a client experiences in therapy (the spike), followed by a gradual reduction in symptoms. These patterns again have been identified in CBT, and studies have shown that non-linear patterns found in the treatment of anxiety are predictive of positive outcome (Heimberg & Becker, 2002).

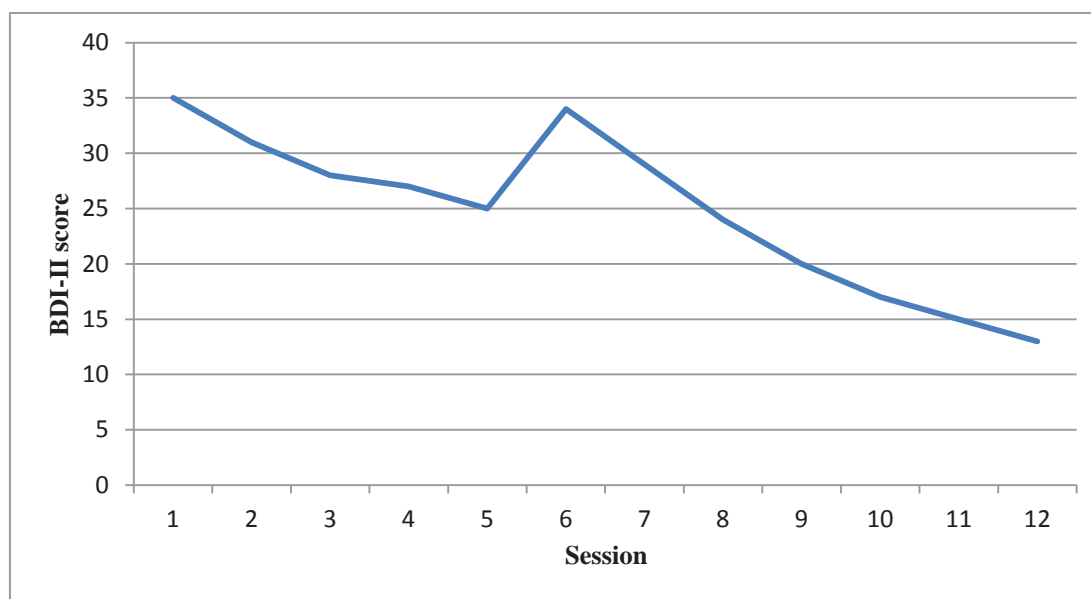


Figure 3.3. Example of a depression spike using Beck Depression Inventory scores.

Client Factors and Change

So far, the research discusses change in terms of how treatment impacts on an individual, reflecting how classic outcome research views change. However, research has

begun to identify associated factors that can impact on how change progresses (Doss, 2004). While earlier studies typically focus on the management of a single problem, it is now clear that in the typical clinical setting, clients show considerable variance in characteristics including distress, problem severity, chronicity, and expectations of improvement (Lueger et al., 2000). Whatever the characteristics may be, they impact on how an individual responds to therapy, therefore hindering the definition of a typical response pattern for any therapeutic approach (Lueger et al., 2000). Leon and colleagues (1999) established a number of significant variables considered to be influential and important in understanding the length of treatment and type of change (based on the change curves shape) that would be expected from a particular client (Lutz et al., 1999). The first of these is subjective wellbeing. Assessed at the beginning of therapy, subjective wellbeing encompasses a number of factors including the client's problem severity, level of distress, issues in everyday functioning, life satisfaction, and emotional adjustment, amongst others. Findings from studies have shown that higher distress at intake results in a higher rate of symptom improvement over the course of therapy (Lueger et al., 2001; Stulz et al., 2007; Thompson et al., 1995). This is thought to be due to a larger amount of change on outcome measures being required before a client with higher distress at intake meets for a clinically significant outcome.

Frequency of symptoms, current life functioning, and experience with past therapy are also viewed as influential. Frequency of symptoms are rated in terms of a specific time period, such as within the past week, month or since the problems began to occur (Lutz et al., 1999). Again, evidence shows that the more frequently symptoms are occurring at the time of presentation to therapy, the higher the rate of change that is likely to occur, although an explanation as to why this occur has not being established (Lueger et al., 2001). Somewhat related to frequency of symptoms, acute, chronic, and

characterological symptoms have been shown to have different rates of improvement over the course of therapy (Kopta, Howard, Lowry, & Beutler, 1994). Acute distress symptoms, characterised by intense emotionality had a more increased responding rate than chronic distress symptoms, which reflected on durable emotional traits, and characterological symptoms, defined as further ingrained symptom patterns than the other two types (Kopta et al., 1994).

Current life functioning concerns the extent a clients' problems interfere with their daily life. Areas to acknowledge include relationships, finances, work, social involvement, and performance in completing routine tasks (Lutz et al., 1999). Although this information is usually provided by clients and from collateral information, current life functioning is typically understood from the clinician's perspective of what is going on for the client, rather than the clients' direct point of view. The higher the score on a measure that reflects their current life functioning, the more rapid change that is expected and the higher the likelihood that therapy will result in a positive outcome (Lueger et al., 2001). Additionally, it would be expected that clients would demonstrate fluctuations in therapeutic progress pertaining to any occurrence of difficulty or significant improvement in these areas while undertaking therapy (Newnham & Page, 2007).

Previous experience with therapy is also an influential factor on rate of change, and refers to any previous involvement with psychotherapy, counselling, or any other mental health involvement (Lutz et al., 1999). Previous therapy typically indicates slower rates of improvement, although this may be linked with the chronicity of the problem (Lueger et al., 2001). Whether or not any previous experience with therapy was positive or negative also impacts on how therapy will proceed, with those who have had a

negative experience more likely to be resistant to a collaborative partnership (Leon et al., 1999).

The length of time a problem has impacted on an individual has been shown to impact on therapy progression (Lutz et al., 1999). The longer a problem has been affecting an individual's life, the slower the expected change will be (Lueger et al., 2001). This further reinforces the importance of the client's involvement, perception of the therapist, and expectations of the effectiveness of therapy as a whole (Lutz et al., 1999; Ryan, Lynch, Vansteenkiste, & Deci, 2011). If a client's expectations of therapy are high, they are expected to achieve change at a rapid rate (Lueger et al., 2001). In general, the way one expects things to happen shapes the experiences they have and how things are perceived (Greenberg, Constantino, & Bruce, 2006).

Two important client factors highlighted in the literature that impact on change are therapeutic alliance, self-efficacy and motivation.

Therapeutic Alliance. The therapeutic alliance refers to the relationship between the client and clinician and is characterised by warmth, empathy, care, competence and genuine positive regard (Westbrook et al., 2011). While clarification around whether the therapeutic alliance influences improvement or improvement influences the therapeutic alliance, findings from research provide a substantial amount of evidence that therapeutic alliance and improved outcomes are associated (Crits-Christoph et al., 2013). In terms of CBT, Beck (2011) identifies the therapeutic alliance as a core principle of this therapeutic modality. As such, characteristics of the therapeutic alliance can influence the type of change an individual is likely to demonstrate while undertaking psychotherapy.

Self-efficacy. A number of therapies, including CBT, consider self-motivation an essential component. As such, a considerable amount of attention should be designated to this form of motivation within therapy, especially if the client is reluctant to work towards change in therapy (Ryan et al., 2011). In addition, and alongside the importance of client expectations of therapy, a client's expectation regarding their own ability to change, and awareness of their current difficulties, is a significant component of successful therapy (Ryan et al., 2011). These expectancies are seen as a function of self-efficacy beliefs, a concept introduced by Bandura (1997). Self-efficacy is based on self-regulation, with the assumption that if a client does not have any belief in themselves that they can change, therapy is less likely to produce positive outcomes (Bandura, 1997). The current breadth of research has been shown to support this notion, indicating that motivation and optimism further enhance positive results (Westra, Dozois, & Marcus, 2007).

Motivation. Motivation encompasses a number of factors, and variables affecting motivation are among the most consistently studied (Mohr et al., 1990). Research has shown that a client's outcome is heavily dependent on their involvement and commitment to the therapeutic process (Ryan et al., 2011), with a greater correlation to positive outcome, client involvement, and rates of responding (Bohart & Tallman, 2010). Prochaska and DiClemente's (1983) previously mentioned SOC model is helpful in identifying which stage of motivation a client presents with in therapy and provides strategies entailing how the client can become more motivated (Norcross et al., 2011). As previously mentioned, this SOC model entails five stages: motivation; pre-contemplation, contemplation, preparation, action, or maintenance, with those in the latter stages of the model improving at a faster rate than those who present to therapy in the former (Prochaska & DiClemente, 1983).

Cognitive Behavioural Therapy and Change

In understanding how clients change while undergoing CBT, the current study focuses on predictions of change based on mood and behaviour and, as such, specific cognitive change is not a predominant focus of the current section. Fennell and Teasdale (1987) compared symptomatic change, as measured by the BDI, and mood change, as measured by the Visual Analogue Scale given to clients before and after each session. Results suggested that those whose BDI symptoms improved quickly demonstrated a simultaneous similar trend in improving mood, whereas those who showed slight symptom improvements had minimal change in mood. Howard and colleagues (Howard et al., 1993) obtained contrasting information finding that mood and emotional wellbeing improved with or without symptom reduction, with symptom distress being unlikely to improve if mood and emotional wellbeing had demonstrated improvements first. Subjective wellbeing and subjective distress, which can reflect the intensity of mood, have also been thought to improve at a rapid rate, indicating that these are the first areas that show change (Maling, Gurtman, & Howard, 1995; McNeilly & Howard, 1991; Simons, Gordon, Monroe, & Thase, 1995). As can be judged by the length of time between the current study and these studies, this area of research requires further updating.

Reflecting on the theory of CBT specifically, contributors have long indicated that initially focusing on an individual's behaviour and mood, followed up by cognitive intervention, will result in enduring symptom relief (Beck, 2011; Hollon & Beck, 2013). Behaviour is typically the first focus of CBT for depression, to engage the client's interest and attention in therapy (Beck, 2011). Often beginning with behavioural activation or activity scheduling, where an individual is encouraged to participate in enjoyable

activities or tasks that demonstrate mastery, such increases in behaviour are thought to result in both improved mood and thought patterns (Cuijpers, Van Straten, & Warmerdam, 2007; Lejuez, Hopko, LePage, Hopko, & McNeil, 2001). Although mood typically improves, it is thought to still remain within the clinical range due to self-critical automatic thoughts continuing to be present prior to cognitive intervention (Beck, 2011). This is due to mood and emotions being seen as an outcome of cognitive and behavioural intervention (Samoilov & Goldfried, 2000) as opposed to being a causal mechanism for behavioural and cognitive change. In terms of response interrelationships, this can be understood as an emotional response providing an incentive for a certain behaviour, with the reduction in negative affect (achieved by participating in enjoyable activities) serving as reinforcement for another response (continued participation in such activities) (Evans, 1999). This suggests that thoughts and mood often will change only after positive events and consequences are experienced more frequently, thus suggesting that behavioural change occurs prior to mood and overall symptomatic change.

When following a CBT protocol in treating anxiety disorder, therapy typically begins with a focus on alleviating the symptoms of anxiety such as the physiological symptoms and maladaptive thinking maintaining the client's difficulties (Wells, 1997). This can be done through introducing relaxation or distraction techniques, as well as activity monitoring and scheduling. When such symptoms have been managed, behavioural techniques are implemented to support the client in confronting their anxiety-provoking stimuli. This is likely to increase an individual's experience of anxiety until such behavioural techniques have been completed. One could then derive from this perspective that anxiety symptoms are likely to be relieved initially, followed by an increase in participating in previously avoided behaviours, alongside increases of anxiety as one engages in tasks aimed to confront the client's fears. Furthermore, this suggests

that one can expect different trajectories of change from a client who is experiencing depression and a client experiencing anxiety disorder (Evans, 2013).

In terms of change following the cessation of therapy, traditional ABA research provides evidence of maintenance, where learned behaviours are continued in the absence of therapeutic intervention, and generalization, where behaviours are continued in settings and contexts similar to the situations in which such behaviours were learned (Evans, 2013) following therapy. In terms of mental health syndromes, maintenance and generalisation can be difficult based on the wide breadth of situations and experiences that a client is likely to be involved with following the cessation of therapy, reflecting a well-documented phenomenon in the ABA literature where an extinguished behaviour is likely to return due to the passage of time as well as following the introduction of novel situations relevant to the behaviour (Evans, 2013). As such, cognitive behaviourists developed a method to train clients to be their own “therapists” by supporting clients in developing and practicing both behavioural and cognitive strategies that can be applied to a variety of situations (Beck, 2011). As such, while behavioural theorists indicate that therapeutic gains may relapse somewhat following the cessation of therapy, a clinician following a cognitive behavioural therapeutic model would expect improvement to be maintained or continued following the cessation of therapy.

Interrelationships Between Overall Symptoms, Mood, and Behaviour Change

When measuring outcomes of CBT treatments, there has been no agreed-upon facet of a client’s functioning that should be measured and reported. The most obvious kind of measure might be the questionnaire that established the client’s syndrome in the first place. Popular questionnaires quickly emerged that were to a large extent symptom checklists. Thus, a syndrome such as social phobia or social anxiety might contain

questions about the intensity of feelings and the nature of avoidance and escape behaviours that tended to occur around social encounters. Measures of anxiety could be in response to a situation (state anxiety) or the more general tendency towards emotional lability and neuroticism (trait anxiety). In the case of depression, the most popular measure of this syndrome was developed by Beck (Beck, Steer, & Brown, 1996), called the Beck Depression Inventory (BDI). Examination of this inventory shows that essentially it is a self-rating of the common symptoms of depression, such as a loss of appetite and disturbances in sleep patterns. Thus the BDI and similarly designed measures of anxiety, such as the Beck Anxiety Inventory (BAI; Beck & Steer, 1990), are standard clinical measurement tools and were used in the present study.

In the original behavioural tradition, however, it was usually assumed that rather than measuring symptomatic change it was necessary to show changes in the actual behaviours related to the general nature of the client's dysfunction, resulting in the introduction of behavioural assessment. For example, the behavioural manifestation of depression is typically a decrease in the number of potentially enjoyable activities that the client engages in during a typical day or week. Thus, a behavioural measure of improvement in the condition known as depression would be a return to everyday functions—the client now doing more and more of the things once enjoyed. Similarly, with respect to anxiety disorders, especially phobias, the most common behavioural measure of improvement was always the willingness to approach the feared object or situation. Sometimes quite simple measures of approach were used, such as the number of centimetres that a client would approach a phobic object, such as a snake. With social anxieties and panic disorder, however, the behavioural indices of improvement were not so easily quantified and a simpler measure could just be the number of activities that had previously been avoided the client was now willing to do.

Of course there are other ways of measuring anxiety and depression, and the most obvious is in terms of what we might loosely call feeling states. In the case of anxiety it could be in terms of activity in the autonomic nervous system, such as rapid heartbeat, sweaty palms, or facial muscle tension. In controlled studies of treatment outcomes these autonomic nervous system changes could be measured directly by psychophysiological recordings, or indirectly by self-report of the sensations. The physiological signs of depression are not so easily measured directly, but are easily estimated by self-report of current mood. Thus, simple self-report measures of negative versus positive mood, such as the Positive and Negative Affect Scale (PANAS; Crawford & Henry, 2004) became widely used and for anxiety the equivalent “mood” measure might be something like the Penn State Worry Questionnaire (PSWQ; Meyer, Miller, Metzger, & Borkovec, 1990).

One of the difficulties that appeared immediately was that these three types of measure (symptoms, mood, and behaviour) did not typically correlate very highly (Evans, 1986). Even within autonomic nervous system measurement of anxiety or sexual arousal, psychophysiological measures and self-report of physical sensations tended not to correlate, nor did they necessarily predict the individual’s behaviour. This gave rise to a long debate within the field as to whether these were simply alternative and inexact measures of a common construct or whether the different types of measure were actually manifestations of different aspects of a psychological construct. A detailed discussion of these questions was presented, without any clear answers, by Evans (1986).

One conclusion that seemed undeniable, however, was that these different aspects of an emotional state were not independent and they influenced each other. For instance an intense feeling of psychophysiological change, such as rapid heart rate, might be judged by a person as sufficiently unpleasant that they would then strive to avoid the

situations that generated such feelings. However if the tendency to avoid and escape could be reduced, regardless of the unpleasant affect being experienced, then the client might experience other benefits, say social rewards, that would in turn help to alleviate the symptoms of his or her syndrome.

At about the same time, a controversy was emerging within applied behaviour analysis along similar lines. When child clients, or people with developmental disabilities, were referred for treatment, they often exhibited more than one problematic behaviour. A child with conduct disorder, for example, might be aggressive (pull hair and kick peers), noncompliant with instructions, and impulsively break social rules (running away, destroying property). In the standard applied behaviour analysis tradition each of these behaviours would be independently targeted, or, even more commonly, only the most serious or annoying behaviour would be identified for change. A number of researchers, but most notably Meyer (Voeltz) and Evans (Evans, Meyer, Kurkjian, & Kishi, 1998; Voeltz & Evans, 1982), recognized that these supposedly independent behaviours were all actually part of an internal system in which one behaviour might easily be the trigger for another. Staats (1968) had described this as one behaviour within a repertoire being the independent variable for another behaviour within the person's repertoire—behaviours, feelings, and symptoms, are all interconnected so that changing one dimension might facilitate, hasten, or cause changes in the other.

While these ideas are implicit within CBT, it is more usual for outcome research to focus only on one dimension of change, most often the symptom pattern as indicated by the BDI or the BAI. But to understand the process of change it might be necessary to have a theory of how different modalities inter-relate. Would we, for example, expect an index of mood/feeling to change before the individual started to engage in new

behavioural patterns? Or would it be necessary to reduce the intensity of the clinical syndrome before the client's self-reported mood/feelings became more positive? As these process inter-relationships during treatment have not been studied very directly in the past, but are implicit in cognitive-behavioural theory, I decided to provide my participants not just with one measure of clinical change whose course they could anticipate, but three different dimensions of change, each one of them easily enough monitored in real clinical practice: getting estimates of frequency of positive behaviours, getting estimates of mood or feelings, and getting estimates of symptomatology. Interestingly, in real clinical practice one might not want to administer a measure such as the BDI every single week, partly because client's symptoms do not always change much from week to week (although, as I have explained, they can change very rapidly) and partly because these symptom checklists are quite short and one's responses from the previous administration are easily remembered. This is why some of the other questionnaires that were designed to be used regularly, such as Lambert and Finch's (1999) Outcome Questionnaire, might have been more typically used by clinicians interested in monitoring change, but I expected my participants to be even less familiar with these as routine monitoring devices. In any event, what I ended up using was a hypothetical measure of behaviour, of mood/feelings, and symptoms, and whether participants would expect these to change together, in parallel, as equally valid reflections of progress, or whether they would expect there to be inter-connected patterns, was one of the research questions.

In terms of the research literature, it became clear that these questions have not typically been asked, and so there is no clear empirical evidence as to which, if any, pattern is most likely. But there are strong theoretical expectations. In behaviour theory approach and avoidance behaviours are considered to be motivated by emotion (two-factor theory of emotion; Schachter & Singer, 1962). Thus, behaviourists following this

logic would expect a change in mood/feeling to result in the overcoming of avoided and inhibited behaviours, and once behaviours were taking place, the benefits would be seen in general symptom reduction. Traditional psychopathology researchers would argue that the syndrome must be alleviated before the client would experience benefits in behaviour, but that symptoms and mood/feeling might well follow a parallel track. Some approaches to CBT, especially Acceptance and Commitment Therapy (ACT; Hayes, Strosahl, & Wilson, 1999), emphasise the need to make life changes (“get out of your mind and into your life” is one of their mantras) and so an ACT theorist might anticipate behaviours changing first, resulting in a reduction in the core syndrome (symptoms of psychopathology), which in turn would result in improved mood/feelings. It can be seen from these possibilities that actually any sequence has some sort of plausible basis, and so the interest of this study is - would participants have a theory about sequence of dimensions of change, or not, and if so, which sequence was most often expected?

Summary

Following on from theories of change, it is becoming clearer that client change is no longer viewed as solely smooth and linear, especially when understanding change in CBT. It is also becoming clearer that what a client brings to therapy is also likely to have an impact on therapy, specifically with respect to their commitment to change and the development of insight as they progress through therapy. Such factors can impact on the rate of change and may be linked to the non-linear patterns that have been identified in the research thus far. While research is still developing in understanding the causal pathways of CBT, and with very limited research on how overall symptom, mood, and behavioural change interrelate with each other, the treatment manuals imply the differences in trajectory between these factors highlighting differences between what

would be expected of individuals being treated for a mood disorder or an anxiety disorder.

CHAPTER FOUR: THE NATURE OF THE CURRENT RESEARCH

As outcome research currently stands, it is evident that an improvement in understanding is needed. When reflecting on the four features of change, that is, having something that needs to be changed, establishing evidence that literal change has occurred, defining an end point by some measure of success, and identifying the pattern that was followed in achieving successful change, classic outcome studies have typically focused on the second and third of these and have had difficulty in understanding and judging clinically meaningful outcomes in terms of identification and process—what really has to be targeted and by what processes has change occurred. These studies also underemphasise the importance of individual client differences and therefore there is a need to consider other ways of incorporating a scientific methodology, such as monitoring progress as you go and adjusting treatment accordingly.

Reflection on components of research that can enable a better understanding of how clients improve in therapy focus on understanding how clients change session to session, as opposed to the difference between a client's scores on a measure before and after treatment. This is to understand the process of change during treatment and identify similarities and differences that can occur between clients and problem areas. As research becomes more capable of understanding change in this way, the research literature has shown that there are different possible trajectories and patterns of change, with recent findings highlighting that change is not likely to be smooth and linear. While the difficulties have been noted, clinicians are guided by the scientist-practitioner model, and thus, if practicing clinicians are using their judgement of client progress to monitor change and adjust treatment, they need to have some sort of conceptual model of what to

expect in individual clients. Based on this premise, clinicians should then be ready for changes highlighted in the research such as sudden gains, depression spikes, and even deterioration. This then begs the question – what are the *implicit models of change* that clinicians have? Do they expect the aforementioned patterns of change, do they expect change to be linear, or do they, for example, expect it to reflect Howard and colleagues (Howard et al., 1986) often cited decelerating linear curve? Based on the breadth of literature pertaining to the scientist-practitioner gap, it remains an important question to consider in terms of how much conceptual work is being absorbed by practicing clinicians or clinical students in training. Therefore, differences between experienced and trainee clinicians are another factor to be explored. While not previously considered in the literature, it was hypothesised that as trainee clinicians are required to address the literature in their training they would be more likely to make predictions similar to what research has determined when compared with their experienced counterparts. It was expected that clinician's predictions would be more grounded in therapeutic experience and, due to the discussion around the scientist-practitioner gap, it would be expected that such predictions would differ from the research.

Awareness of these issues is important because a therapist's expectation of what should be happening to a client is going to determine the way he or she evaluates client progress, might influence the areas that are emphasized in treatment (the therapeutic targets), and might affect the therapist's perceptions of success and satisfaction with the rate and meaningfulness of the client's progress. In terms of overcoming the scientist-practitioner gap, such information would also provide insight into what is being seen in everyday clinical practice. Therefore, this study aims to see how clinicians, some in training and some experienced, think about the likely process of change and the shape of client improvement. Alongside identifying the form of slope predictions change reflects,

this present study additionally aims to understand whether change is expected to be gradual and steady, or characterised by fluctuations in any of the three aspects of change. This connects to an additional aim, which concerns the extent to which therapists expect discontinuous patterns, a feature of change identified in the research literature that is found particularly in clients who are receiving CBT. As such, the current study explored whether or not such discontinuities were a part of clinicians' implicit expectations of client change.

While the patterns of change expected in client progress is of primary interest in the present study, there is also an interest in the level of change expected at the end of therapy. The current literature indicates that it takes, on average, 14 sessions to "recover" (no longer exhibit the symptoms of the syndrome), with research indicating that clinicians can be too optimistic when predicting positive outcomes for their clients. In terms of what constitutes a statistically meaningful change in research, Jacobson and Truax's method of clinical significance is often used. As such, the present study aims to understand what level of improvement each hypothetical client is expected to achieve by the end of therapy. Furthermore, the current study also aims to understand whether or not treatment gains are maintained at 3 and 6 month follow-up to explore what clinicians expect to happen with clients' symptoms, mood, and behaviour, whether that is to improve, deteriorate, or remain unchanged.

Finally, when understanding change in its different components, for example overall symptom change, behavioural and mood change, what do clinicians expect? Will they expect that a client must improve their mood or anxiety before their behaviour will improve, or will behavioural improvement result in mood or anxiety improvement? From these kinds of studies we can gain insight into possible needs for training and exposure to

more complex ideas about the nature of change and the value of session by session monitoring of clients. In the next chapter I describe the study and how it was conducted.

To address these aims, I designed a study to look at clinical judgement, based on improvements to the task used in my Master's dissertation (Fletcher, 2011) which was limited by inconsistent methods in obtaining data and an absence of statistical analyses. Therefore, the present study involved clinicians making predictions based on two hypothetical cases, a client with depression and a client with anxiety, due to the literature's suggestion that the trajectories of change for depression and anxiety are different. As it is becoming clearer that change cannot always be captured by the sole use of symptom measures, I additionally looked at change in the context of mood and behaviour using three ways to measure such outcomes that were likely to be familiar to participants. Furthermore, this study aimed to use measures that could be incorporated easily into everyday practice and are likely to be familiar with practicing clinicians (despite it having been indicated that routine outcome measurement is difficult to implement), in order to see how successful they anticipate client change to be and whether or not they expect it to be long lasting.

In summary, the current study aimed to address the following research questions:

1. What theories of change are clinicians aware of?
2. Do clinicians expect change to reflect a linear, curvilinear, or some other pattern? Are there differences between experienced and novice clinicians?
3. Do clinicians expect Sudden Gains, Early Rapid Responses, and/or Depression Spikes? Are there differences between experienced and novice clinicians?
4. Overall, is change expected to be smooth across therapy or demonstrate variability?

5. In what order do clinicians expect each client to change when considering overall symptoms, mood and behaviour (i.e. do clinicians expect differences between symptom, mood, and behavioural change or think it is the same thing?).
6. Do clinicians expect clients to achieve clinically significant symptom reduction by the final session of therapy?
7. If clinicians anticipate clients will experience clinically significant symptom reduction, do they expect it to be long lasting?

CHAPTER FIVE: METHOD

The methodology for the present study has been developed and extended from Fletcher's (2011) unpublished Master's thesis. The purpose of this chapter is to provide an overview of the design adopted by this research to achieve the aims of the thesis. Ethical considerations are reviewed first, followed by methodology, an overview of participants, a review of the instruments used to acquire data and procedures used. Finally, how data was analysed is highlighted.

Ethics

As this research involves human participants, Massey University Human Ethics Committee (MUHEC) was approached for approval regarding a Low Risk Notification. The decision to apply for a low risk notification was made following the completion of a screening tool which indicated minimal risk to participants. Clients were fully informed prior to participating in the task, provided with appropriate avenues to indicate consent and deception was not deemed to be of concern for this study. Experience in a previous study of a similar nature also provided insight into the low levels of risk this research posed. MUHEC thus approved the present study, and permission to proceed with this research was granted.

Methodological Framework

The current study proceeded with a qualitative research design using session-by-session measurement as the guiding methodology. This was based on the benefits of using session-by-session measurement to understand change as highlighted by patient-focussed research (Howard et al., 1996). Presented in the literature review above, patient-profiling

(Lambert et al., 2001a; Lambert et al., 2003; Lueger et al., 2001), feedback research (Hawkins et al., 2004; Lambert et al., 2001b; Lambert et al., 2002b) and research regarding discontinuous patterns (Hayes et al., 2007b) have also utilised this. Taking this into account and noting the main aim of the study was to understand change over the course of therapy, it was deemed appropriate that the expectations of change were predicted for each session of therapy as opposed to specifically pre- and post-measures.

When presenting clinical cases in research, three methods have been highlighted. The first involves writing about a client candidly and directly following informed consent (Sperry & Pies, 2010), the second entails presenting clinical cases while disguising the clients' characteristics, and the third involves case composites, where a number of real clients characteristics are combined into a hypothetical single case, providing another option in writing about clients in research. Each method has their advantages and disadvantages, with completely undisguised case studies offering direct experience of a clinician and their client, while potentially compromising or altering the therapeutic relationship. As such removing identifying factors may remove changes in the relationship that may arise from obtaining consent (Sperry & Pies, 2010). However, one problem that can arise from removing identifying factors concerns achieving a balance between changing the clients' characteristics enough to protect their identity and ensuring the case represents the client's presentation enough to uphold validity (Patterson, 1999). Although completely hypothetical cases allow control over client characteristics by ensuring that cases are not too complex while also capturing realistic situations that do not impinge on client confidentiality, and thus aiding overall objectivity (Grundy, 2008), constructing real life is not real life. Therefore, a disadvantage of using a hypothetical case study reflects the inability to verify the data (Aas & Alexanderson, 2012).

Acknowledging the advantages and the disadvantages for each of the above methods, for the purpose of the present study, case composites have been used to allow the development of two cases that are simple enough to not overwhelm participants with too many characteristics to consider, as well as being similar to clients that may present in the clinical environment. In this study in particular, an undisguised case study was not deemed appropriate as the focus was on how a particular client was *expected* to change not exactly *how* a specific client changed. A case study on its own has potential to differ from a typical client whereas a case composite can collate the similarities between cases. In addition, case studies are typically used in understanding the treatment and experiences of the client being discussed, whereas the purpose of this study focuses on how characteristics influence predictions, and thus the inclusion of a real client is not warranted. As indicated by Lincoln and Guba (2002), ideal case studies should be creative and outline an independent and personal construction of the data being observed, a principle which is reflected in this study's hypothetical cases.

Using a hypothetical case to collect data on expected session-by-session change reflects a quantitative design with clinician experience reflecting the dependent variable and expected outcomes reflecting the independent variables. Further insight into the research questions was provided from qualitative information provided by open ended questions.

Participants

Two groups of participants were sought for the present study. One group consisted of experienced clinicians, with the other group being students involved in a clinical psychology programme. Different approaches were used to obtain participants for each group between August, 2011, and November, 2012. Experienced clinicians were those

who identified themselves as being registered as a clinical psychologist. Clinical psychology students were considered potential participants after the completion of their first year at practicum level through one of the clinical psychology programmes offered within New Zealand.

Measures

The electronic task, presented in Appendix A, was a development and extension of the task used in the study undertaken by Fletcher (2011). The task describes two hypothetical case scenarios. The first described “Mr T”, a client with depression, and the second described “Ms S”, a client with anxiety. These hypothetical case descriptions, explained the client’s current living and vocational status, the symptoms which they were exhibiting, the extent of their disruption to functioning, and treatment goals. Also included were the initial scores for three measures; one based on mood, one based on symptoms, and one based on behaviour. For the client exhibiting depressive symptoms, this entailed the Beck Depression Inventory, the Negative Affect schedule of the Positive and Negative Affect Schedule, and a collaboratively devised measure based on participation in activities. For the client demonstrating anxiety symptoms, the measures used were the Beck Anxiety Inventory, the Penn State Worry Questionnaire, and a goal driven self-report diary reflecting the percentage of regular activities she managed to complete each day. For each measure, a description of its purpose was provided alongside a short guide for the interpretation of scores. With the sessions controlled for 12 weekly sessions, and two follow-up sessions (at three and six months), participants were then asked to predict the scores for the following 13 sessions on each of the measures for each client on a specialised graph. This graph included three Y axes to account for the highest score that could be achieved for each measure. By doing this, each measure could be

plotted against the other, thus allowing for participants to demonstrate, or adjust their expected change patterns, based on their potential assumptions regarding the differences between mood, symptomatic, and behavioural change.

One of the limitations of the previous study was the use of an electronic task for experienced clinicians while trainee clinicians completed a paper tasks (Fletcher, 2011). The electronic task involved participants entering scores manually on a graph on a computer screen, rather than manually plotting the graphs in a paper task. In the present study, the electronic task was used by both clinicians and trainees.

The use of relevant and familiar measuring tools was important in these case studies to improve the expectation that the prediction of change was based on an understanding of the construct being measured. A number of recommendations have been made with regards to measures in outcome research. Bickman and colleagues (2000) suggest that a measure that is individualised, easily provided, and responsive to the construct being measured is appropriate. In addition, it is also recommended to use measures that are able to be repeatedly administered (Barkham et al., 1998).

When identifying which outcome measures are most often used in the clinical setting there is no consensus (Barkham et al., 1998). A study administered to clinicians worldwide found that although there was a large variety of measures being used, the Symptom Checklist 90 Revised (SC;-90-R), the HRSA, the Beck Anxiety Inventory (BAI), the State Trait Anxiety Inventory (STAI), and the Beck Depression Inventory (BDI) were most often used (Society for Psychotherapy Society for Psychotherapy Research, 1993). Similar findings were found in a study by Froyd (1996) who identified the BDI, the STAI, and the HRSD as being most used among his participants, with Mellor-Clark and colleagues (1997) further finding the BDI, the SCL-90-R, BSI, and the

Inventory of Interpersonal Problems being the most commonly used in the United Kingdom. A more thorough understanding of the measures used by clinicians and an understanding of how clinicians use the results will provide much more comprehensive information. Although the feedback literature provides a number of methods to assess outcome, implementation of these methods is not possible in the current study as data must be accrued over time from the same population, and assume a linear progression of change.

For the purpose of this study, and through consultation with an experienced clinician, the following six measures were chosen for the two case studies presented. In seeking the type of change that occurs within the confines of the therapeutic process, it was noted that three different outcome measures would present a comprehensive understanding of how change is expected to occur. These measures reflect: 1) conventional psychometric assessment of symptoms, 2) client affect such as negative mood in depression and worrying in anxiety, and 3) client behaviour. Based on a previous dissertation (Fletcher, 2011), it is expected that change will reflect improvement on the first two measures, with behaviour expected to improve over the duration of therapy. Through the use of two case studies, the client goals are established as the relief of symptoms, to feel better, and to participate in regular activities that have been neglected since the onset of the client's problems. These goals reflect the three change types.

'Mr T.'

The hypothetical case description of Mr T provided symptoms supportive of a diagnosis of clinical depression including low mood, loss of appetite, poor sleep, inability to concentrate, alongside it becoming increasingly difficult to find things enjoyable especially in terms of activities he previously liked doing. Mr T indicated that his

treatment goals were to “experience less negative moods on a day-to-day basis, to have a healthier, happier lifestyle (including the reduction of depressive symptoms), and to be able to get out and do more enjoyable things with his family or by himself, the way he used to do in the past”. The treatment protocol set in this case is outlined as follows.

CBT for depression. Cognitive Behavioural Therapy (CBT) has been a highly researched therapy used in the treatment of a number of psychological problems, including depression (Friedman & Thase, 2007). To ensure a level of control over the general course of CBT the client was likely to undertake, Mr T’s case scenario provided a description of the hypothetical treatment protocol. In doing so, the process of therapy illustrated in this scenario entails a focus on current problems and building goals that the client is realistically able to work towards.

The described protocol indicates a standard evidence-based CBT plan for depression restricted to 12 sessions and 3 and 6 month follow-up based on insurance limitations. This treatment plan included the use of psycho-education, cognitive restructuring, homework assignments, the teaching of coping skills and encouragement, along with the instillation of hope in the client to participate in activities.

Additionally, cues were provided regarding factors that have been empirically shown to impact on the course of change – motivation and insight (Duncan, 2010). This involved a change in motivation in session four, and the development of increased insight in session six. In the development stages of this thesis, a research aim was developed to reflect whether or not clinicians acknowledged an influence of such factors and how it affected the predicted rate of change, however, preliminary analyses indicated that few clinicians indicated a significant change following the indicated sessions and thus in-depth analyses were not carried out. Three types of outcome measurement were used to

show the participants' implicit ideas on change in the therapeutic process for this case, the Negative Affect Schedule, Beck Depression Inventory and a record of positive activities. These measures were thought to be familiar psychometric methods for participants and appropriate alongside a CBT treatment protocol based on research and anecdotal evidence.

1) Beck Depression Inventory. The Beck Depression Inventory (Beck, Ward, Mendelson, Mock, & Erbaugh, 1961) consists of 21 items in which an individual is asked to choose one statement out of four options that best represents aspects of their functioning during the week prior to administration. Rating for each statement ranges from low endorsement of depressive symptoms to high endorsement.

The BDI was developed using clinical consensus based on the observation of typical attitudes/symptoms of depressed clients, and the summarisation of these. Following a need for clarification after the initial development, the BDI was modified to produce the BDI-II which is the current form used in clinical practice and research. The BDI-II has largely appealed to clinicians due to its ability to identify depression in a timely and cost effective way. The BDI-II's use is also backed by the wide breadth of research available for this measure (Groth-Marnat, 2003). Furthermore, this measure has been frequently and successfully used in the assessment of both clinical and non-clinical individuals (Groth-Marnat, 2003).

Being the preferred choice in numerous studies as well as clinical practice, the BDI-II has been subjected to extensive psychometric evaluation (Dozois, Dobson, & Ahnberg, 1998). This inventory has demonstrated acceptable reliability, with an internal consistency typically between .89 to .94 and test-retest reliability approximately at .93 following a one week interval (Beck et al., 1996). Furthermore, the BDI-II has been

shown to have adequate concurrent validity and the ability to discriminate between clinical and non-clinical populations (Groth-Marnat, 2003).

The BDI-II's utility in the hypothetical case of Mr T was to measure the clinical symptoms of depression and how they evolve over therapy. Assessed with the BDI-II on initial consult, the individual in this case received a score of 35, indicating severe depression.

2) Negative affect schedule. One half of the Positive and Negative Affect Schedule (PANAS; Watson, Clark, & Tellegen, 1988), the Negative Affect Schedule (PANAS-NA) involves the rating from one to five, with one being *very slightly/not at all* and five being *extremely*. These items were based on five mood domains; distress (distressed, upset), anger (hostile, irritable), fear (scared, afraid), guilt (ashamed, guilty), and jitters (nervous, jittery), and clients rated them based on the extent they experienced these moods in a clinician specified time period (this moment, past few days, past few weeks, past year etc.) (Watson et al., 1988). For the purpose of this study, the time period was indicated as one week.

The PANAS as a whole measure has demonstrated excellent convergent validity (Crawford & Henry, 2004; Watson et al., 1988) and has also shown adequate reliability (Crawford & Henry, 2004) with the Cronbach's alpha for the negative affect schedule alone being $\alpha=0.85$ (95% CI = 84 – 87). When assessing short term periods of distress, the PANAS has demonstrated sensitivity in identifying mood fluctuations, and demonstrates trait stability in long term distress assessments.

Mr T's initial assessment provided a score PANAS-NA of 26, deemed a high level of negative affect when compared to the average adult score of 14 (Nezu et al., 2000).

3) Self-monitoring of positive activities. To understand behavioural change for Mr T, self-monitoring was implemented as homework. In therapies influenced by behavioural theory, the identification and monitoring of the frequency of concerning behaviours are an important part of following change in the therapeutic process (Whipple & Lambert, 2011). Furthermore, this emphasis on self-monitoring is based on the goal oriented nature of CBT (Kuyken, Watkins, & Beck, 2007) with the behaviour being monitored relating to one of the therapeutic goals. It allows the goal to not only be addressed within therapy, but also outside therapy, while also providing both client and clinician with a way to identify whether or not improvements are being made (Kazantzis, Whittington, & Dattilio, 2010). For the case of Mr T, self-monitoring addressed the goal of increasing involvement in activities that he had once found enjoyable, but at the initiation of therapy was unable to find pleasure when participating in them.

While not subject to psychometric evaluation, self-monitoring has been shown to be a useful method of assessment, as well as a therapeutic technique. The key to the success of behavioural self-monitoring has been conceptualised as being linked to both compliance and motivation (Burns & Spangler, 2000). In association with this use of self-monitoring, the participants receive a hint that Mr T becomes more motivated at a specific stage during treatment.

“Ms S.”

Ms S is described as having symptoms of a diagnosis of generalised anxiety disorder (GAD). These symptoms consist of panic/agoraphobic-like symptoms including heart palpitations, feeling dizzy, a feeling that she was about to pass out, sweaty hands, and hyperventilation. Her symptoms were described as being most prominent when

driving, and she also had high levels of discomfort when at the mall or a crowded supermarket. As such, Ms S. had also been avoiding numerous “day to day” activities.

The goals Ms S. set for herself included to become free of the symptoms of GAD, to reduce worrying and her constant feelings of nervousness, and to enable herself by engaging in ordinary everyday social activities that she had begun to avoid prior to the initial assessment.

CBT for Generalised Anxiety Disorder. Being the most common psychological problem that help is sought for, it is unsurprising that the treatment of anxiety disorders has been subject to extensive research, especially in terms of the preferred treatment strategies (Schneier, Mellman, & Spiegel, 2007). Currently, CBT has been the treatment modality shown to be more effective than pharmacological interventions when targeting Panic Disorder and GAD (Schneier et al., 2007). Therefore, the treatment overview provided in the case study for Ms S followed a 12 session general CBT oriented protocol. Like Mr T, the number of sessions were controlled by articulated health insurance restrictions. The protocol expressed to the participants of this study included psycho-education, specifically on the nature of fear and anxiety, exploring the relationship between fear and avoidance behaviour, teaching of deep muscle relaxation and calming breaths, setting a hierarchy of activities and proceeding through it, as well as exploring catastrophic thinking. In keeping with the present study’s aim in understanding the impact of motivation and insight cues on the course of change, participants in this study were provided with subtle descriptions of situations pertaining to these two concepts. Again, hints were provided to participants regarding factors that can influence change in therapy including a change in motivation in session four, and the development of increased insight in session six. As preliminary analyses suggested that clinicians did not

demonstrate any perturbations in the sessions following, an aim pertaining to how these factors affected predictions of the rate of change were not included in the current study.

Three outcome measurement methods were used to hypothetically assess client change over the course of therapy and at three and six month follow-up. These were the Beck Anxiety Inventory, the Penn State Worry Questionnaire, and self-report of the percentage of daily activities completed per week. These methods were thought to be appropriate and of some familiarity to most participants, although a description of the first two inventories is provided to participants, along with the scores associated with varying levels of functioning.

1) Beck Anxiety Inventory. Appropriate for individuals aged over 17 years of age, the Beck Anxiety Inventory (BAI) (Beck, Epstein, Brown, & Steer, 1988) entails 21 descriptive statements reflective of anxiety symptoms. Reflecting on the past week, individuals are asked to choose one of four ratings describing the extent to which this symptom has been having an impact on their functioning (Campbell-Sills & Brown, 2010).

The BAI was developed based on a need to devise a measure that discriminated between depression and anxiety (Turk & Wolanin, 2006). It has been shown to have excellent reliability with a Cronbach's alpha of 0.92 and a test-retest reliability of .75 (Beck et al., 1988). The BAI has also demonstrated satisfactory concurrent validity, construct validity, and discriminant validity. The central reason as to why the BAI is an often used measure is based on its ability to not overlap with depression measures based on its focus on fears and hyper-arousal symptoms (Campbell-Sills & Brown, 2010). While superior in assessing panic symptoms, the BAI is a commonly used measure in the assessment of anxiety (Turk & Wolanin, 2006).

In the present study, the purpose of the BAI was to keep track of the client's anxiety symptoms between each session. In the hypothetical case of Ms S, participants were informed that she scored 48, which fell within the "severe anxiety" range (Campbell-Sills & Brown, 2010).

2) Penn State Worry Questionnaire. The Penn State Worry Questionnaire (PSWQ) is comprised of 16 statements narrowed down from 161 derived from clinical and research findings, daily diaries from GAD patients, and items that had been used in an already developed cognitive somatic anxiety based inventory (Startup & Erickson, 2006). The 16 items relate to different situations in which worry can occur and individuals rate them based on how they would react. This is done by rating each statement from one to five, with one being not at all typical of me, and five being very typical of me (Startup & Erickson, 2006).

Used to evaluate current level of worry, the PSWQ (Meyer et al., 1990) was developed when GAD progressed from a residual diagnostic category in the DSM-III-R (American Psychological Association, 1987). To better understand the concept of worry central to a GAD diagnosis, the development of a psychometric tool was required for clinical use (Startup & Erickson, 2006). Along with the aforementioned avenues that were used to identify the initial 161 items, the existing theoretical views of worry also influenced the development of this measure.

According to Startup and Erickson (2006) the PSWQ's popularity is due to this measure being the "most widely used measure of the frequency, intensity, and uncontrollability of worry (p. 101)". Furthermore, it has demonstrated high internal consistency in both clinical and non-clinical groups exhibiting a Cronbach's alpha between 0.88 and 0.95. It also has good test-retest reliability, $r = 0.74-0.92$ and has

established sensitivity to change across both six weeks and 12 weeks (Startup & Erickson, 2006). The PSWQ has often been used in the case of applied research

The purpose of the PSWQ in the case of Ms S is to measure her worrying during the week between sessions. At initial consult, her score is stated to participants as 73, which is considerably high in comparison to a score of 32 which indicates mild/rare worrying.

3) Self-monitoring of the percentage of daily activities completed. Due to the importance that self-monitoring has in CBT, as suggested in the description of outcome measures for Mr T, it is again implicated in the case of Ms S. While Mr T's self-monitoring focuses on the completion of 10 enjoyable activities, Ms S focuses on the percentage of appropriate daily activities she participated in between sessions. The use of self-monitoring in clients with anxiety disorders enables clinicians to follow changes based on a common coping strategy utilised in anxious individuals, avoidance. By establishing that an individual is increasing the amount of time they are spending in situations that typically result in heightened anxiety, it indicates that they are reducing their participation in avoidance strategies (Wells, 2006). Thus, such a behavioural measure on specific target behaviour provides "real information" about the client's change progression (Newman, Ciarlo, & Carpenter, 1999).

As previously mentioned, there are no psychometric properties that have been derived for self-monitoring, nor is there an abundant amount of information regarding its use as an outcome measure in anxiety disorders. However, it has been supported as a tool that is often utilised in CBT in the monitoring of client progress (Hopko, Armento, Cantu, Chambers, & Lejuez, 2003; Whipple & Lambert, 2011).

For the case of Ms S, behaviour change was described as being monitored through completion of a daily events diary. Following the hypothetical development of a list in the initial session consisting of reasonable activities that Ms S would like to participate in between sessions, overall behaviour for the week was predicted based on the average percent of tasks Ms S was able to complete since the prior session. In addition, it was suggested that the number of activities Ms S would be likely to participate in would increase week to week as therapy progressed. As a baseline, participating clinicians were informed that Ms S's initial level of participation was 0%.

Additional Questions

For the purpose of gaining demographic information to provide further insight into results derived from this study, a series of questions were posed. Participants were asked whether or not they are a student with an answer of yes leading to a question regarding how many years they have participated in the clinical programme. For those who indicated that they were not a student, a question regarding the number of years of clinical experience was posed, along with their professional job title. A question regarding familiarity with the measures was also posed to ensure that lack of familiarity with the chosen measures did not create a limitation. Participants were then asked if they had experience with the treatment of depression and anxiety with CBT. The experience question was posed separately for each client type. This was to account for differences and similarities that may be found between individuals who practice within different theoretical orientations.

Based on the limitations found in Fletcher's (2011) study, where the task information alone provided minimal understanding of the influences on predictions, participants were asked if they were aware of any models or theories based on the

prediction of client change. If they selected yes, participants were then asked to name or describe the theories they were aware of. Additionally, participants were directly asked to reflect on the process of their predictions of client change and outline if there were any ideas, concepts, or theories that they could think of that impacted on their expectations. Participants were also given the opportunity to make any further comments they thought were fitting.

Data Collection

Experienced clinicians were invited to participate through a number of avenues. Firstly, Massey University faculty were approached, based on their association to the clinical psychology programmes. Contact was also made to the heads/administrators of the New Zealand Council of Clinical Psychology (NZCCP) and the New Zealand Psychological Society (NZPsS), with an invitation distributed through their email lists. Following a lack of participants from the initial invitation, the psychological advisors of each of New Zealand's DHBs were approached, and an email sent to the clinical psychologists within their respective areas. Personal connections of both researcher and supervisor were also used.

In seeking student participants, permission was first sought from the Director or Head of Department of these programmes. The programmes approached included those based on the three campuses of Massey University: Albany, Turitea, and Wellington, as well as the University of Auckland, Victoria University, Canterbury University, Otago University, and the University of Waikato. Once permission was obtained, the administrators of each of these departments forwarded the link for the electronic task to each of their students. The invitation provided by NZCCP and NZPsS would also have been sent to student subscribers.

Invitations were scheduled to be sent out towards the end of 2011 to ensure those in the clinical programme had experienced their first year which typically includes a form of practicum. Follow-up reminders were again sent in February, 2012 to students, as well as the NZCCP resending the invitation to their members via email in May, 2012.

The emailed task began with an information sheet depicting the general area which this research aimed to study, what criteria needed to be met to be deemed eligible to participate, and the conditions of consent. Consent was granted by participants submitting their task which they had the option of doing once completed. Aside from demographic details, participants did not supply any personal information. Following the completion of the task, they were provided with the option of submitting their name and email address to enter the draw to win a voucher. This information was kept separate from their task to ensure participant anonymity.

Analyses

To analyse the aims proposed by this study the Statistical Package for the Social Sciences (SPSS) version 21 (IBM Corp, Released 2012) was used, with each aim resulting in a different method of analysis, as discussed below. Prior to addressing the aims, the data were coded to ensure that experienced and trainee clinicians were differentiated to allow for comparison throughout the analyses.

Preliminary assumption checks

Prior to beginning analyses the assumptions of normality were checked to ensure that the statistical strategies used were appropriate and to ensure that interpretations of the findings were accurate. Following the guidelines provided by Field (2009) to check the assumption of normality, that is, that the distribution of data is approximately normal, a

normal distribution was plotted on top of the histograms and visually inspected. Furthermore, “Normal Probability Plots” (P-P plot) were used to plot the residuals for each measure predicted to ensure that they formed a diagonal line, thus confirming a normal distribution (Field, 2009; Singer & Willett, 2003).

Analysis 1: Clinician awareness of change theories

To understand what theories of change clinicians were aware of, the data were coded by each theory that was indicated by participants. As this was an open question, participants were able to identify any theory they associated to change and were not solely related to the theories discussed in the literature review. From this, the frequencies of each possible change theory were identified and compared between experience levels.

Analysis 2: Shape of change over sessions 1-12

As covered in the review of the literature, the collection of session-by-session data has become increasingly preferred in studying change over the course of therapy, as opposed to the comparison of pre- and post-treatment scores. As such, attention has focused on statistical analyses that can effectively address such data collection. One method that has received attention is Multilevel Modelling (MLM), with this method being applied to the aim of understanding the pattern of change measure for symptom, mood/anxiety, or behavioural change for both Mr T and Ms S.

Multilevel modelling (MLM)

MLM is a regression based statistical method comparable to Analysis of Variance (ANOVA) and ordinary least squares regression (OLS) with one predominant difference: MLM accounts for “nested” data (Kahn & Schneider, 2013). Nested data refers to hierarchical data collected over a number of levels, for example, individuals receiving

treatment *grouped* by the therapist they were treated by. MLM then analyzes this nested data with a two level approach to change. Level 1 describes within-person change allowing for the consideration of individual change trajectories and thus allowing for the understanding of how individuals differ from each other in their slopes and intercepts. With reference to the above example, this would be in terms of the difference in change trajectories between each client being treated. Level 2 involves between-person change which considers the effect of particular variables on the change trajectories across a number of individuals, that is, how individuals can be grouped to understand how grouping/nesting factors influence the variability in intercept and slope. Again referring to the above example, this could mean addressing how trajectories of clients treated by therapist X differ from the trajectories of clients treated by therapist Y. This is the simplest form of the MLM approach to nested data, with the ability to address differences between further groupings, for example, if the above example wanted to identify the difference between change trajectories of clients treated by therapists in clinic A versus clients treated by therapists in clinic B. MLM is thus an effective method for investigating individual change, as well as considering the effects of different levels of data collection.

While MLM allows for the analysis of either level, most approaches to MLM collapse Level 1 and 2 models together algebraically into a simple composite statistical model (Graham, Singer, & Willett, 2009). For Level 1, growth curve modelling is typically applied to explore or test hypotheses regarding individual variability in change trajectories. In this case, it is assumed that intercepts (i.e. initial change point) and slopes (i.e. shape of the change trajectory) have fixed and random components, termed fixed effects and random effects. Fixed effects entails modelling change trajectories using an average intercept and slope across all participants. Random effects represent the degree to

which the intercept and slope vary between individuals (Level 1) and across individuals in different groups (Level 2).

As such, MLM analyses typically involve developing a series of models, applicable to the research question, that address the fixed intercepts and slopes where an average intercept and slope are modelled for the average change trajectory, and account for random slopes and random intercepts to determine the degree to which intercepts and slopes differ between participants. These two effects can be addressed individually or simultaneously, thus making MLM a “mixed model” approach. This then allows researchers to address which one of these models best fits the sample data being analysed, as well as addressing the degree of variability that may be associated between individuals. Furthermore, MLM allows for researchers to address what type of slope best reflects the trajectories of change, whether that is a linear trend, a quadratic trend, a cubic trend, or some other predetermined nonlinear trend.

Level 2 factors are analyzed in MLM by applying time-variant and time-invariant factors. Time-invariant factors are factors that remain the same over time, for example, gender typically remains constant over time. Time-variant factors are factors that change as time progresses, for example, client motivation in therapy. MLM can thus utilise such factors to identify participants who represent different levels, such as participants who identify as male and female, and identify how this impacts on the variability between slopes and intercepts (Kahn & Schneider, 2013).

While the analysis strategies ANOVA and OLS have been long used as the statistical approaches often used for addressing change within therapy, there are a number of advantages to using MLM. The primary advantage of MLM is that it is able to address two or more levels of change using nested data as indicated above. In contrast, ANOVA

and OLS expect change to be equal among participants and thus treat within-group variability as error. As such, if researchers are interested in understanding if change trajectories differ between individuals, ANOVA and OLS are not appropriate (Kahn & Schneider, 2013). Furthermore, MLM's incorporation of nested data has been shown to reduce the occurrence of Type 1 error (Field, 2009). In addition, while ANOVA assumes that variances and covariances are equal over time, MLM can handle data with different variance-covariance structures. It can also analyze linear and nonlinear patterns (Hedeker, 2004).

MLM is further thought to be advantageous due to its flexibility concerning how data is collected. MLM allows for data to be collected at either a set interval or at varying time points for each individual without affecting the validity of the models developed. Validity is also upheld when the number of data points collected varies for each participant. As such, incomplete data can be analysed alongside complete data (Hedeker, 2004).

Applying MLM to the current study, data have been nested by session, by clinician, and by experience level. As such, analyses involved growth curve modelling to address individual differences for session-by-session predictions of change by participating therapists reflecting analysis of Level 1 change. Furthermore, MLM was used to address Level 2 change by addressing similarities within each experience level, out of experienced and training, as a time-invariant factor. As such, the current study is able to use MLM's advantages associated with nested data.

For change to be modelled using MLM, Singer and Willett (Singer & Willett, 2003) stated that three important features must be considered. Firstly, three or more waves of data are required, any less and the shape of each participant's individual

trajectory cannot be described and true change and measurement error cannot be distinguished. Eleven waves of data were analyzed (Sessions 2-12). This involved all sessions in which treatment/therapy was administered. The baseline score (Session 1) was excluded as this predetermined score was not a clinician response. The decision to exclude the two follow-up sessions (3- and 6-month follow-up) in this analysis was that no intervention occurred during this phase of treatment. As such, analyzing sessions 2 to 12 provides a homogenous data in which there was active intervention. In addition, while MLM allows for such discontinuities in time, the focus of including follow-up sessions was to evaluate change following therapy as a separate specific aim, thus addressing a separate issue to the pattern of change. As such, these sessions will be addressed through separate analyses.

Singer and Willett's (2003) second requirement is that a sensible metric of time is used. The present study treated time as weekly sessions, thus reflecting equally spaced time points. Session numbers were reverse coded so that session values ranged from -10 to 0, making the final session (Session 12) the intercept for the analysis. Reversing sessions was thought to be advantageous in the current analyses as identifying the average level of functioning at the end of therapy was deemed more valuable based on the present study's interest in the final outcome of therapy. While MLM allows for missing data, clinicians were not able to leave scores blank and thus each participant made a prediction for each time point.

Finally, Singer and Willett (2003) indicated that change research should be evaluated using a relevant, continuous outcome measure that demonstrates systematic change over time. Psychometrically sound measures are essential in capturing change empirically. For the present study, measures were chosen based on their relevance and

adequate psychometric properties for the construct being measured for both the mood and symptomatic measures. While psychometric properties are not available for the behavioural measures utilised in this study, the research literature provides evidence for the use of behavioural self-report in the assessment of client change (Michie & Johnston, 2012).

When performing MLM analysis, three main assumptions must be met; linearity, normality and homoscedasticity (Singer & Willett, 2003). As previously stated, the assumption of normality was confirmed for the presented data via visual inspection of histograms, kurtosis statistics and inspection of P-Plots. Linearity, where it is assumed that the relationship being modelled is linear, was addressed through scatter plots of the standardized residuals against standardized predictors (Field, 2013). Linearity was confirmed when the scatter did not appear to have a curve in the residuals. Homoscedasticity, the assumption that the variance of the residual terms for each outcome variable, was confirmed if the scatter did not appear to funnel and was evaluated as being roughly equal widths apart (Field, 2013).

Having had the assumptions of MLM met, the covariance structure of relevant data should be defined. If the covariance structure applied is too simple then the probability of a Type I error is increased. In contrast, if the covariance structure identified is too complex, a Type II error becomes more probable. When running MLM with repeated measures, first order autoregressive structure is typically applied (Field, 2009; Singer & Willett, 2003). This covariance structure assumes correlations across repeated measurements are higher at adjacent time points.

The MLM procedure was applied to the current analysis to understand whether or not a model of the expected trajectories of change is improved when taking into account

the structure of the data as well as variability between slopes and intercepts (Heck, Thomas, & Tabata, 2010; Singer & Willett, 2003). As such, increasingly complex models were built up that took into account random effects (i.e. variability in slope and intercept) and data structure as well as increasingly higher order polynomial functions in order to determine the shape of change over the course of a 12 session CBT protocol.

Model 1 and Model 2. Singer and Willett (2003) recommend the development of two models prior to using MLM. The first (Model 1; unconditional means model) involves the examination of random effects which identifies how much of the variance in predictions of each measure's expected outcomes (i.e. the level-1 variable) is attributable to within-person variance (i.e. differences between clinicians; level-2 variable) using the intraclass correlation (ICC). The ICC is a correlation coefficient used to determine the relative magnitude of the between-person variance components (Singer & Willett, 2003). It is determined using the within-person variance estimate of the intercept σ_{ϵ}^2 and the between-person residual variability σ_0^2 where $p = \frac{\sigma_0^2}{(\sigma_0^2 + \sigma_{\epsilon}^2)}$. The ICC value provides how much variance is accounted for by between-group variance, with a noteworthy ICC indicating that considering level 2 factors may be worthwhile, thus acting as confirmation for the use of MLM for the modelling of time-variant and time invariant factors (Singer & Willett, 2003).

Model 1 is then followed by an unconditional fixed linear growth model (Model 2) to address how much of the within-person variance is accounted for by time, in this study's case, by session. In Model 2 time is treated as a predictor for level-1 variance, assuming that Session will have a linear effect on all predictions. Resulting significant fixed parameter estimates of the intercept and slope values suggest that it is reasonable to assume clinicians report different intercepts *and* linear slopes of improvement, thus

providing rationale for the development of a third model allowing for the modelling of random slopes. If both random slopes and intercepts improve model fit (Model 2), these effects were then incorporated in all subsequent models.

While the unconditional means model (Model 1) and the unconditional linear growth model (Model 2) provide confirmatory information for the appropriateness of MLM, they also provide a source of comparison for subsequent models by comparing residual variance, discussed later in this section.

Model 3. Model 3, the random intercepts and slopes linear growth model, involved introducing random effects allowing clinicians' individual intercept and rates of change to vary while accounting for the structure of the data. As previously mentioned, the covariance structure typically assumed for repeated measures data is the first order autoregressive and, as such, this structure was applied to Model 3.

Model 4: The current study was also interested in whether or not change reflected a quadratic trend consistent with Howard and colleagues' (1986) description of a decelerating curvilinear trend of change. As such, Model 4 (the random intercepts and slopes quadratic growth model) represented Model 3 with the addition of a quadratic effect for Session. This effect was created by squaring time (Session²) and adding it as a fixed and random effect.

Model 5. Hayes and colleagues (Hayes et al., 2007a) further indicated a possible cubic relationship for change, accounting for depression spikes and other possible discontinuities that reflect change that changes direction twice, thus creating a curve with a double bend. In MLM a cubic trend is modelled by multiplying time by itself three times (Session³). As such, the random intercept and slopes cubic growth model (Model 5)

was developed by adding the cubic effect of session to Model 4. When building models using MLM, higher order polynomials are typically added until there is no substantive improvement in model fit determined by significant changes in residual variance between models.

Model 6, 7 and 8. Once a model that best captures the pattern of change across therapy had been determined, analysis focussed on identifying if there were any group differences in change over the course of therapy by entering the dichotomous variable Experience (Experienced clinician vs. Trainee). As clinician experience is a time-invariant factor in the current analyses, this effect was addressed by adding experience as a fixed effect to the model for each measure that achieved a better fit. As such, providing that the previously discussed models demonstrated improved fit each time, Model 6 would model the effect of experience on a linear growth model, Model 7 would model the effect of experience on a linear and quadratic growth model, and Model 8 would model the effect of experience on a linear, quadratic, and cubic growth model.

Additional features of MLM. Two components are of interest when using MLM: fixed effects and random effects (identified as variance components in the analysis). Fixed effects address the systematic interindividual variability the trajectories predicted, in accordance to the values in the level-2 predictor, providing an overall equation that represents the average intercept and slope of change for all clinicians (Singer & Willett, 2003). In the current study fixed effects are evaluated by the five parameters: γ_{00} , γ_{01} , γ_{10} , γ_{20} , γ_{30} . Parameter γ_{00} reflects the average true initial status (i.e. intercept) based on the grand mean of all participants when session equals 12. Representing the average participant γ_{10} provides an estimate of the associated rate of change (i.e. the true slope) for a linear model. When identifying if a quadratic or cubic model fits the data better than

a linear trend, γ_{10} represents the average instantaneous rate of change when session equals 0 and γ_{20} represents the overall average quadratic effect of time. Similarly when considering a cubic fit, γ_{30} represents the additional average cubic effect of time. In the multilevel equation, γ_{01} is defined as the difference between the true initial status and the intercept reflective of individuals within subgroups, for example, the difference between trainee and experienced clinician predictions, with γ_{11} representing the difference between individuals in different subgroups in their rate of change.

Recoding sessions from 0 to -11 to -11 to 0 and running the analysis again allows for the determination of the magnitude of change at the end of therapy to compare with the magnitude of change at the beginning of therapy (i.e. γ_{00}). This analysis was run again using the recoded variables to allow for the model to identify the average score and average variance at session 12 which are two useful measures of treatment impact.

Variance components describe how much participants vary from fixed parameters by providing their own set of parameters (Singer & Willett, 2003). For a Level-1 model, σ_{ϵ}^2 represents an estimate of the within-person variance for all participants to identify how much predictions differ from γ_{00} . Level-2 parameters include σ_0^2 , which represents the between-person residual variability for the predicted intercept, and σ_1^2 , which reflects the between-person residual variability for the predicted rate of change (slope) when evaluating linear effects. When evaluating the between-person effect of a quadratic trend on the rate of change, the value σ_2^2 is used, with σ_3^2 used for cubic trends. The covariance between the true intercept and the true slope across individuals is demonstrated by σ_{01} .

MLM identifies how well the developed models fit using the Pseudo R^2 and Goodness-of-fit Statistics (Singer & Willett, 2003). The Pseudo R^2 represents the amount

of unexplained outcome variation that can be explained by the addition of explanatory variables into a model through the equation:

$$\text{Pseudo } R^2 = \frac{\text{residual variance}_{\text{fewer}} - \text{residual variance}_{\text{more}}}{\text{residual variance}_{\text{fewer}}}$$

Significant decreases in the residual variance between each model indicate that the newer model accounts for more of the unexplained variance than the prior model, thus indicating that the additional factors have a substantial impact on the model's fit. Of note, and in contrast to traditional R^2 statistics, the Pseudo R^2 can result in a negative outcome which Singer and Willett caution as largely uninterpretable (2003).

Goodness of fit in MLM is typically measured by a chi-square likelihood ratio test, the -2 log-likelihood (-2LL), which compares observed frequencies with those predicted by the model (Field, 2009; Singer & Willett, 2003). Although -2LL is often used and is the initial log-likelihood ratio provided, two adjustments, Akaike's information criterion (AIC), and Schwarz's Bayesian criterion (BIC) are often reported. The AIC and BIC both take into account the number of parameters estimated in the model with BIC penalizing complex models more, making it best used when sample sizes are large and the number of parameters is small (Field, 2009). For all three statistics, the smaller the value, the better fitting the model is. When comparing models, the goodness of fit can be addressed by assessing the change in -2LL by subtracting the new likelihood ratio from the old likelihood ratio. To evaluate whether a model is a better fit than the prior one, the three goodness of fit statistics (-2LL, AIC, and BIC) were produced for each model. Due to higher order polynomials increasing the probability of overfitting the data, BIC was the preferred likelihood ration as it penalizes model complexity more than -2LL and AIC (Field, 2009). To simplify this discussion by utilising only the relevant

deviance statistics to this study, only the -2LL and BIC are discussed in the results section. To compare models the change in -2LL was determined using the equation $-2LL_{\text{Old Model}} - -2LL_{\text{New Model}}$. The resulting value was then compared to the critical chi-square value related to the change in degrees of freedom between the two models ($df_{\text{Old Model}} - df_{\text{New Model}}$). If the difference between the -2LL values for each model exceeded the critical chi-square, the model is deemed as having a better fit (Field, 2009). Similarly, change in BIC was determined using the equation $BIC_{\text{Old Model}} - BIC_{\text{New Model}}$. Although not a significance test, a better fitting model was determined by Rafferty's (1993) indication that a BIC difference exceeding 10 is substantial enough to confirm a better fit.

Although MLM provides an analysis of aggregate change (fixed effects) and allows insight into variance across clinicians (random effects) to a certain extent, at times there may still be a substantial amount of variance that cannot be explained or adequately described. As such, individual growth curves were analysed further using a standard OLS multiple regression analysis to test the fit of linear, quadratic and cubic models for each individual clinician. Using Field (2009) George and Mallery's (2014) guidelines, this entailed first creating two new variables for time squared (Session^2), and time cubed (Session^3) to reflect a quadratic and cubic trend respectively. Using the original Session variable as the linear trend, these variables were then entered as models to be tested simultaneously using forced entry option to sequentially test linear, quadratic and cubic models of fit. A 'best fit' model was identified, taking into account both explained variance and also model complexity, and when R^2 was significant ($p < .05$) as judged by the change in F-ratio. A higher order polynomial model was determined as better fitting than the prior model only if R^2 was significant and the increase in R^2 between models was

more than .02 (IBM SPSS Analytic Catalyst, n.d). Visual inspection of the curves was also used to confirm and validate the best fitting model. A new variable was created following the identification of the best fitting model for each clinician's predictions with a non-significant fit coded as 0, a linear fit coded as 1, a quadratic fit coded as 2, and a cubic fit coded as 3.

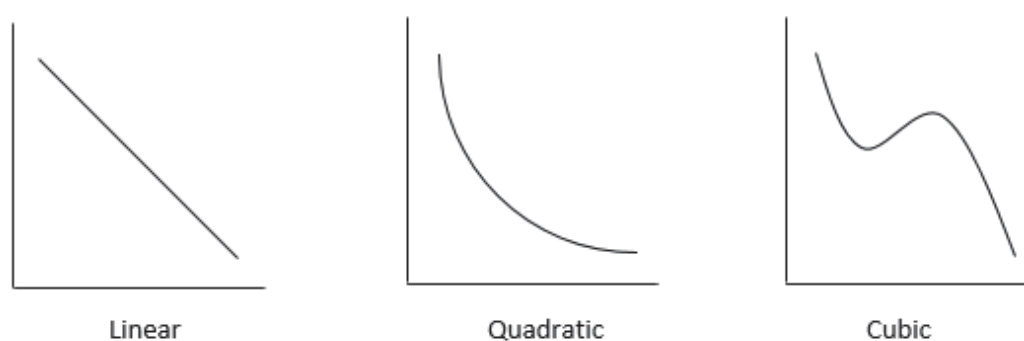


Figure 5.1. Examples of linear, quadratic, and cubic curves.

Using this new variable, a chi square analysis was used to identify the percentage of the occurrence of each 'best fit' model.

Analysis 3: Presence of discontinuous patterns of change

To analyse clinician predictions of early rapid responses (Ilardi & Craighead, 1994), sudden gains (Tang & DeRubeis, 1999b), and depression spikes (Hayes et al., 2007a), methods to identify such patterns were identified in the literature. While the original studies typically involved the use of the BDI-II, studies evaluating sudden gains in individuals with anxiety symptoms have used the 7-point improvement based on the BAI being designed to parallel the BDI (Beck & Steer, 1990; Ingram, Overbey, & Fortier, 2001; Present et al., 2008). Likewise with Stiles' (2003) findings, a 7-point improvement on the BAI is close to an RCI of 1.96, indicating reliable change. Following the application of these methods to identify the presence of non-linear patterns, a new

variable for each Mr T and Ms S were computed to identify the frequency that each pattern was predicted by clinicians.

Analyses addressing how many clinicians predicted either Mr T or Ms S are likely to experience an early rapid response involved Ilardi and Craighead's (1994) definition where 60-70% of the client's overall change is achieved by session 4, with overall change being calculated between session 1 and session 12. To simplify analysis, the lower range (60%) of Ilardi and Craighead's (1994) definition was used to encompass all participants who would fulfil this criterion.

To identify if clinicians expected clients to experience a sudden gain at some point over the course of therapy, Tang and DeRubeis' (1999b) definition was used, where three criteria must be met (using the BDI in the equations): Firstly, the change must be large in absolute terms, with Tang and DeRubeis operationalising this as a decrease of 7 points on the BDI-II or BAI ($BDI N - BDI N+1 \geq 7$). This was determined to be appropriate due to prior reports of BDI score changes reflecting secondary peaks in frequency distribution plots along with Stiles and colleagues (2003) highlighting that a 7 point improvement was linked with the BDI's RCI based on Jacobson and Truax's (1991) method.

Secondly, the gain should be large in terms of symptom severity prior to the gain. This was operationalised by Tang and DeRubeis (1999b) as at least a 25% reduction of impairment in the session prior to the gain as identified by scores on a reliable and valid outcome measure ($BDI N - BDI N+1 \geq 0.25 \times BDI N$). While there is debate as to whether this criterion is relevant based on a lack of justification from Tang & DeRubeis, it was deemed appropriate to include to reflect the current breadth of research on sudden gains, as well as prior removal of this criterion in research showing no differences in the individuals identified as experiencing a sudden gain (Hardy et al., 2005).

Finally, Tang and DeRubeis (1999b) indicated that the three sessions ($N-2$, $N-1$ and N) preceding the gain must have a mean level of functioning significantly lower than the three sessions following the gain ($N+1$, $N+2$ and $N+3$) to reflect the gain being large in relation to fluctuations in depression severity before and after the gain. By this definition, it is difficult to assess the presence of sudden gains early in therapy. A modification that has been used to account for earlier sudden gains while still indicating the change to be large in relation to fluctuations involves the gain being equal to or greater than 1.5 times the individual's standard deviation, across all time points (Kelly, Roberts, & Bottonari, 2007; Kelly, Roberts, & Ciesla, 2005).

Depression spikes were analysed using Hayes and colleagues (Hayes et al., 2007a) definition, where Tang and DeRubeis' (1999b) 7-point change on the BDI-II or BAI was used in the direction of deterioration. Unlike the sudden gain, the 7-point deterioration was not required to be sustained.

Analysis 4: Continuous change versus discontinuous change

Whether or not change was characterised by being smooth and continuous or variable and discontinuous across the course of therapy was also of interest to the present study to identify whether clinicians predicted a series of ups and downs or, rather, change generally followed a smooth progression. A chi-square analysis was then applied to the outcome for each measure to identify whether gradual and steady or variable change was predominant. This thus provides us with additional understanding of what is expected with client change. As an example Figure 5-2 and Figure 5-3 on page 109 demonstrate a smooth and continuous trend and a discontinuous trend, respectively.

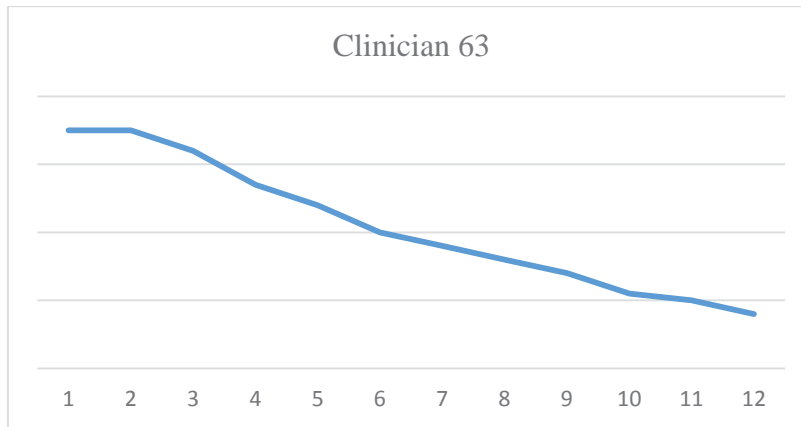


Figure 5.2. Example of predictions reflecting a smooth and continuous predicted trend.

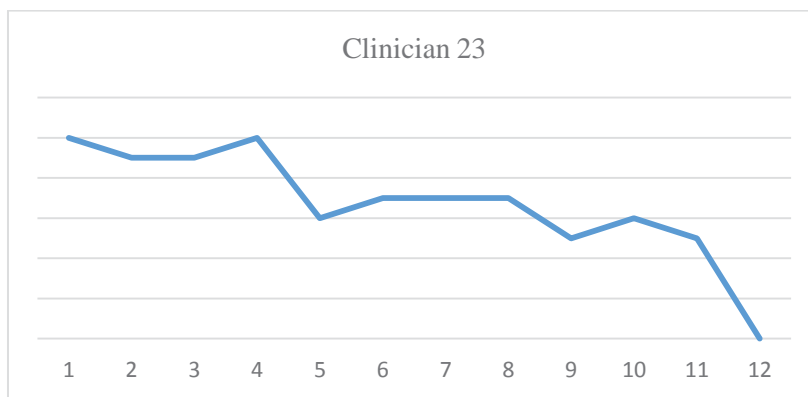


Figure 5.3. Example of predictions reflecting a discontinuous predicted trend.

Analysis 5: The relationship between overall symptom, mood, and behavioural change

The exploration of how overall symptoms, mood and behavioural change interrelate with each other firstly involved visual evaluation of the graphs produced from the electronic task. These graphs demonstrated decreasing overall symptom and mood scores, alongside improving behavioural scores. Using this to visually evaluate the graphs thus created difficulties in assessing how the three types of change interrelated due to the lines of improvement not progressing in the same direction. As the behavioural measure progressed in the opposite direction to the mood and symptom measures, the scores were re-calculated to reflect a decreasing slope as opposed to an increasing slope so that

improvement was represented by decreasing scores, consistent with other measures. This allowed client improvement to be equally reflected in all three measures as a progressively downward sloping graph. For the client with depression, this was done by subtracting the clinician predicted score from 10. For the client with anxiety, this was done by subtracting the clinician predicted score from 100. To then convert the graph to demonstrate all three measures, the overall symptom, mood and behaviour scores were first adjusted to fit on a graph with a single Y-axis. The purpose for this was to produce an easier pattern to compare the three forms of change. For the client with depression, both the BDI-II scores and the PANAS-NA scores were measured on an axis that went up to 50, with number of activities participated in going up to 10. As such, the activity scores were multiplied by 5 to demonstrate plot points reflective of a 50 point axis.

For the client with anxiety, the BAI and PSWQ scores were adjusted to match the maximum 100 the percentage of activities measure allowed for. As the BAI axis allowed for a score up to 60, predicted scores were divided by 60, then multiplied by 100 to produce an equivalent score. Similarly, the PSWQ axis allowed for a score up to 80 and, as such, predicted scores were divided by 80 and multiplied by 100 to produce an equivalent score.

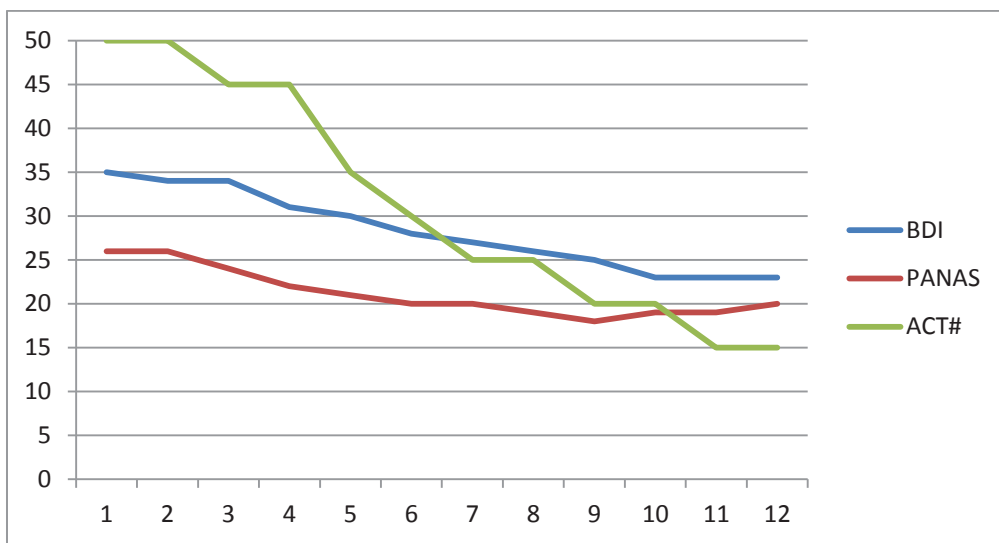
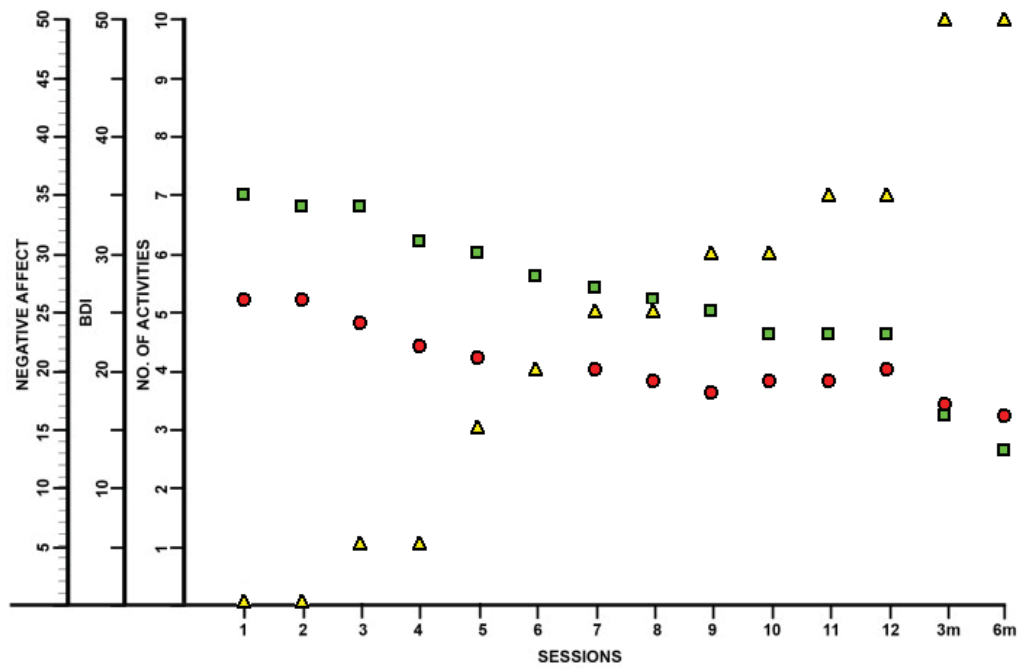


Figure 5.4. Original graph and graph adjusted to show scores progressing in the same direction for client with depression.

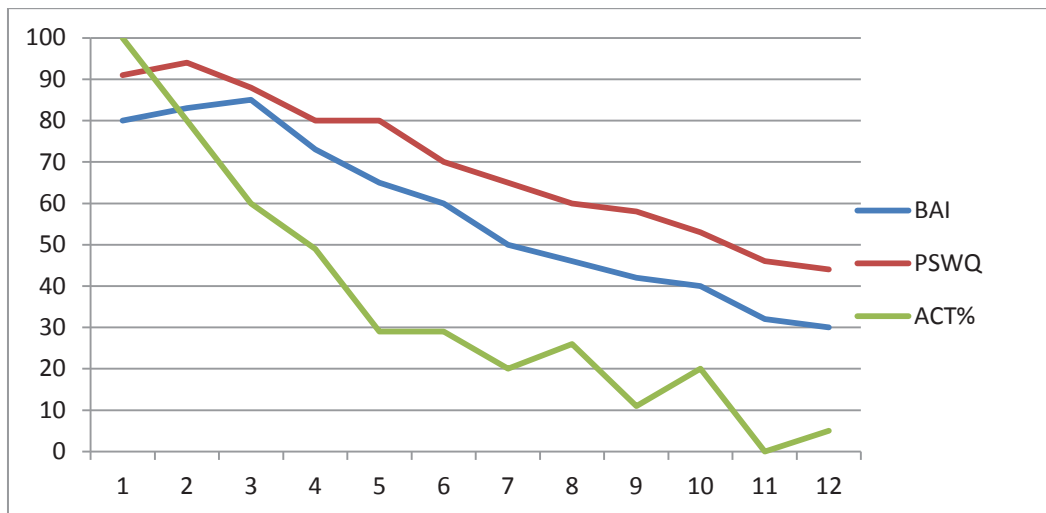
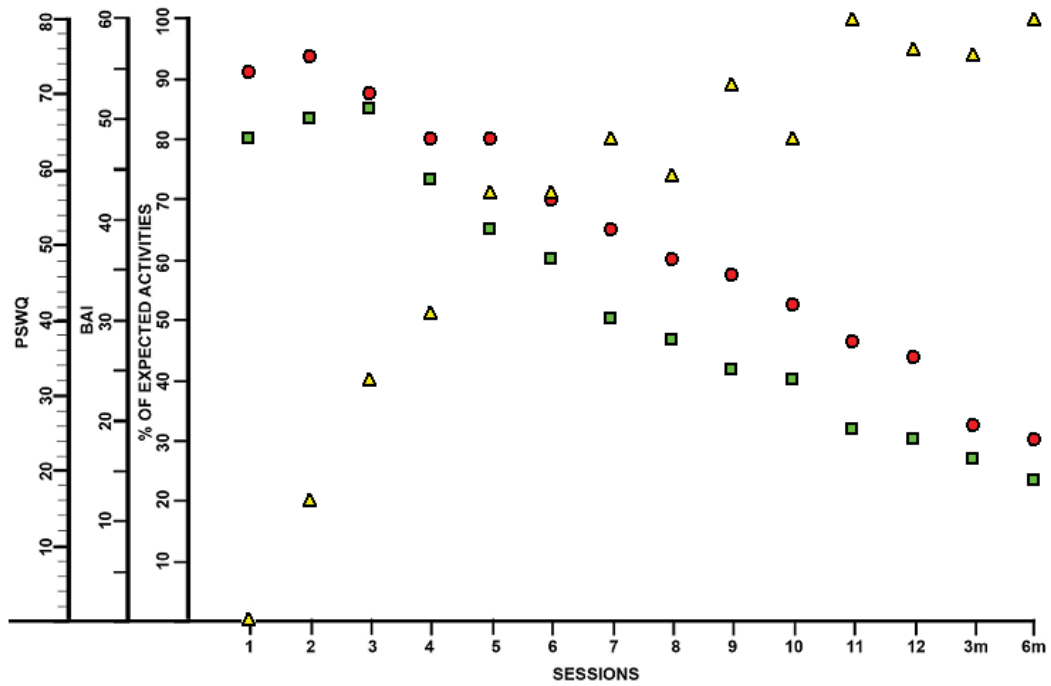


Figure 5.5. Original graph and graph adjusted to show scores progressing in the same direction for client with anxiety.

To further support visual evaluation, criteria were developed to detect an appropriate discontinuity on each measure to identify whether an improvement on one measure preceded or followed improvement on another measure. As previously stated, no obtainable research has been developed to address this form of analysis and, as such, the criterion was developed to identify a change between sessions that exceeded a certain amount. To identify such a substantive change for overall symptoms and mood, the highest score within the non-clinical range for each of BDI-II, PANAS-NA, BAI, and PSWQ, was subtracted from Session 1 provided in the electronic task for each measure. The outcome of this calculation was then divided by 11 (the number of sessions following session 1). This average amount of change required per session for a client to demonstrate non-clinical scores on each measure was then doubled to ensure adequate criteria to demonstrate a reasonable amount of change between sessions. As there is no psychometric data available for the behavioural measures, visual evaluation indicated that a change of 1 for number of activities participated in for the client with depression, and a change of 10 for the percent of activities completed for the client with anxiety, demonstrated a reasonable amount of change. From the identification of amounts that indicated a large, reasonable amount of change, a new variable was created to identify between session changes that exceeded these for each between session change (i.e. between session 1 and session 2, between session 2 and session 3, etc.), with 0 indicating this double average not being exceeded and 1 indicating the double average being exceeded.

Using this new variable for both clients, the case summaries for each measure were placed side by side for simple identification and were considered alongside the

graphs. This allowed an understanding of which type of change occurred first, and in what order changes occurred.

Analysis 6: Clinically significant symptom reduction by session 12

In understanding how many clinicians expect clients to achieve clinically significant symptom reduction by the final session of therapy, Jacobson and Truax's (1991) clinical significance method was used. As discussed in Chapter One, clinical significance is thought to be better at assessing clinical change than statistical significance, with Jacobson and Truax's method being frequently used in psychological research (Crits-Christoph et al., 2013). The analysis focused on the mood and symptom measure scores, PANAS-NA and BDI-II for Mr T and PSWQ and BAI for Ms S, with behavioural measures being excluded due to norms being unavailable. Analyses began with identifying the most appropriate way of addressing the first of two criteria proposed by Jacobson and Truax, the cut-off score. This involves selecting one of three possible cut-offs; A, B or C, presented below, where clinical and nonclinical refer to clinical and nonclinical population norms respectively. Both clinical and nonclinical norms were identified for the PANAS-NA, BDI-II, PSWQ, and BAI.

$$\text{Cut-off A} = M_{\text{clinical}} - 2SD_{\text{clinical}}$$

$$\text{Cut-off B} = M_{\text{nonclinical}} + 2SD_{\text{nonclinical}}$$

$$\text{Cut-off C} = \frac{[(SD_{\text{clinical}} \times M_{\text{nonclinical}}) + (SD_{\text{nonclinical}} \times M_{\text{clinical}})]}{SD_{\text{clinical}} + SD_{\text{nonclinical}}}$$

When applying each cut-off to this study's data, Cut-off A was deemed not appropriate as two of the measures, the BDI and the BAI, result in a score below 0 when taking the sum of two clinical population standard deviations from the clinical population

mean, which is not a possible outcome for either of these measures. Cut-off B was also deemed inappropriate as Jacobson and Truax indicated that this cut-off should only be calculated if nonclinical data is the only data available. As indicated, clinical data is available for all four measures used. As such, with both clinical and nonclinical data available for all four measures utilised in this study, Cut-off C was identified as the most appropriate cut-off for the current study's data, reflecting the weighted midpoint between the means of a functional and dysfunctional population. Using the above equation for Cut-off C, a value was calculated indicating the cut-off for clinically significant change for each measure. A variable was then developed in SPSS for the outcome score on each measure at session 12 to identify those who surpassed this cut-off, thus meeting criterion one for Jacobson and Truax's clinical significance method, and those who did not.

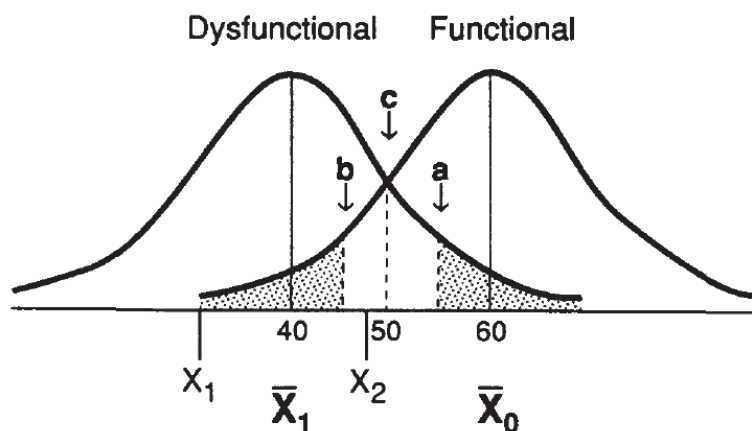


Figure 5.6. Pre- and post-test scores and cut-off a, cut-off b, and cut-off c indications for participant x (Jacobson & Truax, 1991; p14).

The second criterion of Jacobson and Truax's clinical significance method, the Reliable Change Index (RCI) was then calculated. This involved calculating the standard error of measurement (SEM) and standard error of difference (S_{diff}) for each measure, using the following equations:

$$SEM = SD_{clinical\ sample} \sqrt{(1 - r_{xx})}$$

$$S_{diff} = \sqrt{(2 * SEM^2)}$$

As the clients used were hypothetical, and as the first session score on each measure was provided, a standard deviation could not be calculated for the initial session and thus could not be used. As such, and in alignment with Jacobson and Truax's suggestions, the standard deviations from clinical populations for each measure were used. Having calculated the SEM and S_{diff} , the RCI statistic was then calculated:

$$RCI = \frac{X_{post} - X_{pre}}{S_{diff}}$$

The outcome of this for each measure was then used to produce a variable indicating whether or not an outcome score at session 12 met the criteria for reliable change.

Combining the two coded variables to these two criteria, clinical significance classifications were used to identify what kind of change clinicians expected by session 12 where passing both the cut-off and RCI criteria indicated a client who has *Recovered*, passing the RCI criterion but not the cut-off indicates *Improved*, and passing neither criteria indicates *Unchanged*. Using the outcome of this analysis, a new variable was developed for each measure indicating which classification a participant's predicted score met at session 12 and the frequencies calculated.

While the behavioural measure did not meet the criteria to be analysed using the clinical significance method due to the absence of norms or reliability statistics, comments will be made on the average score for session 12.

Analysis 7: Maintenance of therapeutic change at 3- and 6-month follow-up

Following on from the above aim, the current study sought to understand whether clinicians predicted clinically significant change to continue following the cessation of treatment. To understand this, a repeated-measures analysis of variance (ANOVA) was used. This method is suitable when there are two or more repeated measurements of the same people, such as a repeated measure across the course of therapy as used in the current study. This involved the analysis of the scores predicted at session 12, 3 month follow-up and 6 month follow-up on the PANAS-NA, BDI-II, and number of activities for Mr T, and the scores predicted on the PSWQ, BAI, and percent of activities for Ms S. The repeated measures ANOVA was performed separately on each measure. A significant outcome of the repeated measures ANOVA resulted in the individual analysis of change from session 12 to 3 month follow-up, and 3 month to 6 month follow-up using post-hoc pairwise comparisons.

In addition to the assumption of normality, ANOVA also requires confirmation for the assumption of homogeneity of variance. To address the homogeneity of variance, where the spread of scores is roughly equal at different points of the predictor variable, the Levene's test was used. A significant score ($p < .05$) indicates a violation of this assumption, with all six measures used in the current analysis resulting in Levene's test statistics that were not significant. The assumption of sphericity, that is, the assumption that differences between all pairs of groups have equal variances, is also important when completing an ANOVA with repeated measures. While it is not essential for sphericity to be assumed to complete a repeated-measures ANOVA, corrections are required should this assumption be violated (Field, 2009). This assumption is tested by Mauchly's test of sphericity with a non-significant statistic ($p > .05$) indicating the assumption of sphericity

being met. If Mauchly's test of sphericity is significant ($p < .05$), it is recommended that the Greenhouse-Geisser correction is used if Mauchly's test statistic is under 0.75 to avoid a high number of false null hypotheses failing to be rejected (Field, 2009). If Mauchly's test statistic is over 0.75 then it is recommend that the Huynh-Feldt correction should be used (Field, 2009).

An additional complementary analysis determined the frequency with which clinicians reported improvement, no change, or deterioration. Based on this, a variable was computed to identify the amount of change that occurred between these sessions, alongside a variable to identify whether a clinician predicted that Mr T or Ms S's scores would deteriorate (coded -1), demonstrate no change (coded 0), or improve (coded 1) for each between-session change. Using the two new variables, the frequency of clinicians who predicted Mr T and Ms S to deteriorate, show no change, or improve was identified. In addition, the magnitude of change was determined to reflect on whether this improvement was substantial using the Reliable Change Index (RCI) introduced in the prior analyses. Changes in clinical significance classification were also coded as deteriorated, no change, or improvement.

CHAPTER SIX: RESULTS

The results section first presents an understanding of the demographic characteristics of the research sample. This is then followed by the results of each aim in the order they were discussed in the method section. Following the first aim, findings are presented individually for Mr T and Ms S to provide a coherent understanding of both cases.

Demographics

A total of 68 participants accepted the invitation to participate in the present study. Of these, 34 (50%) were experienced clinicians and 34 (50%) were trainee clinicians. Deriving the same number of clinicians for each experience level was achieved by chance. Of the experienced clinicians, 29 (85%) reported their current profession as a clinical psychologist, 2 (6%) stated they were a psychologist, 2 (6%) indicated being a member of academic staff, and 1 (3%) indicated being a psychotherapist. Years of experience ranged from 1 year to 25 years, with the average being 6 years. Of the trainee clinicians who participated, 11 (33%) indicated being in their first year of a clinical psychology programme, 9 (27%) were in their second year, 7 (21%) in their third, one (3%) in their fourth, and 2 (6%) in each of the fifth, sixth and seventh year.

In terms of experience and familiarity with CBT in treating depression and anxiety, 32 (94%) experienced clinicians and 18 (53%) of trainee clinicians indicated prior experience of treating depression with CBT, with 33 (97%) of experienced clinicians and 15 (44%) of trainee clinicians indicating experience with treating anxiety with CBT.

Reflecting on the measures used in the two hypothetical cases presented in this research, 10 (29%) experienced clinicians were familiar with all measures used, 23 (68%) were familiar with some of the measures, and 1 (3%) indicated not being familiar with any of the measures. Out of the trainee clinicians, 8 (24%) were familiar with all measures used, 25 (74%) were familiar with some of the measures, and 1 (3%) was not familiar with any of the measures.

Preliminary Analyses: Assumption Testing

As highlighted in the prior section, a number of assumptions were tested prior to the completion of analyses as per statistical method requirements. The current section presents the results for the assumption testing on PANAS-NA predictions. First, normality was addressed by a Normality P-Plot. Figure 6.1, below, demonstrates the confirmation of the assumption of normality for the PANAS-NA measure, as the residuals form an approximate diagonal line showing limited deviance from normality.

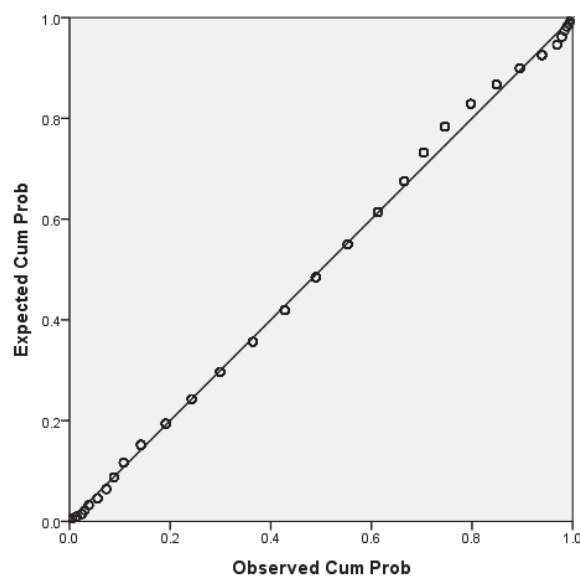


Figure 6.1. Residual plot for negative affect score (PANAS-NA scores).

For analyses requiring an analysis of variance (ANOVA), the assumption of homoscedasticity was required. In addition, for the MLM analysis, the assumptions of linearity required confirmation. To address these two assumptions, the scatter plots of the standardized residuals against standardized predictors were visually evaluated. The PANAS-NA scatter plot is provided in Figure 6.2, below. Demonstrating an approximate rectangular scatter, this plot indicates that the data does not curve, confirming the assumption of linearity, and does appear to funnel, thus confirming heteroscedasticity.

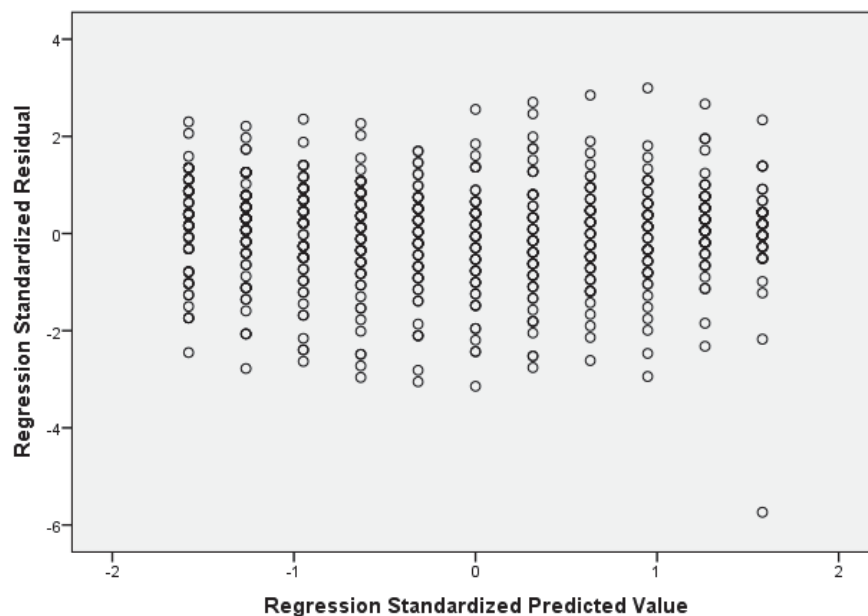


Figure 6.2. Standardised residual scatter plot for negative affect across time (PANAS-NA change).

The assumption of heterogeneity, while not necessary for MLM, is another assumption that required testing before conducting a repeated measures ANOVA. This assumption posits that the spread of scores is roughly equal, and thus can be visually assessed by looking at the standardised residual scatter plot. The Levene's statistic for each measure was calculated for each measure using the SPSS guidelines from Field (2009), resulting in a non-significant result across all measures, thus confirming the assumption of homogeneity.

To address the homogeneity of variance, where the spread of scores is roughly equal at different points of the predictor variable, the Levene's test was used. A significant score ($p < .05$) indicates a violation of this assumption, with all six measures used in the current analysis resulting in Levene's test statistics that were not significant.

While the assumption of sphericity is also required to complete repeated measures ANOVA, this will be discussed in the section regarding this analysis based on its influence on the interpretation of results.

Analysis 1: Clinician Awareness of Change Theories

In asking participants if they were aware of any models or theories related to predicting change in therapy, 78 % (53) of all experienced and training clinicians indicated that they were not aware of any change models. When understanding this within experienced and trainee clinicians, 68% of experienced clinicians indicated that they were not aware of any particular change models compared with 88% of trainee clinicians, with this being a significant difference, $\chi^2 (1) = 4.19, p < 0.05$.

Models of change specified by 32% of experienced clinicians included the Stages of Change Model, with one of these individuals also indicating an awareness of Howard and colleagues (1993) Three Phase Model. One clinician indicated their awareness of Sudden Gains theory, but did not highlight dynamical systems theory as a single theory. Other "models/theories" identified were Scott Miller's model, Cognitive theory, Acceptance and Commitment theory, Ian Evans, Bradford Keaney, as well as Watslawick and Fitch. The Stages of Change Model, the therapeutic alliance, Roc*ROI, and Carl Rogers were the theories of change identified by the 11% of trainee clinicians who indicated their awareness of a model of change.

Mr T**Analysis 2: Shape of Change over Sessions 2 -12****Model A: Overall symptom change**

The parameters for each model developed in the multilevel modelling of the predictions of BDI-II scores are provided in Table 6.1 (pp. 124-125).

Model 1 and 2 represent the unconditional means model and unconditional fixed linear growth model respectively. As time has not been accounted for in Model 1, only γ_{00} is provided, representing the fixed estimate of the intercept. This indicates the average BDI-II score across all participants, estimated as 21.766 ($p < .001$). MLM also provides estimates of the random effects, termed variance components. Variance components at both Level 1 (within-person) and Level 2 (between-person) are addressed by σ_{ε}^2 and σ_0^2 respectively. In Model 1 of the MLM analysis for overall symptom scores, both of these estimates are significant at the $p < .001$ level.

To determine whether Model 1 estimates the level of between-person variance which provides rationale to take into account variability between clinicians using a random linear growth model (i.e. Model 2), the ICC is calculated using the equation $p = \frac{\sigma_0^2}{(\sigma_0^2 + \sigma_{\varepsilon}^2)}$. The ICC value for Model 1 was thus determined to be 0.189 for BDI-II predictions, indicating that 19% of all the variability in BDI-II predictions is between clinicians, suggesting that it may be worthwhile to consider intra-individual differences.

Table 6.1.
Results of Fitting a Taxonomy of Multilevel Models for Change on the BDI-II Across the Course of Therapy

Parameter	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7
Fixed effects							
Initial status, π_{0i}	γ_{00}	21.766*** (0.539)	11.051** (0.571)*	11.051*** (0.772)	12.464*** (0.667)	12.814*** (0.623)	12.650*** (0.936)
Exp	γ_{01}						-0.373 (0.550)
Rate of change, π_{10}	γ_{10}	-2.143*** (0.038)	-2.143*** (0.075)	-1.201*** (0.020)	-0.644* (0.247)	-1.160*** (0.210)	-0.749** (0.268)
Session ²	γ_{20}			0.094*** (0.019)	0.240*** (0.063)	0.094*** (0.191)	0.138*** (0.026)
Session ³	γ_{30}				0.010* (0.004)		
Exp: Session	γ_{11}					-0.082 (0.149)	-0.904* (0.379)
Exp: Session ²	γ_{12}						-0.087* (0.037)
Variance components							
Level-1	Within-person	σ_{ϵ}^2	61.060*** (14.213)	10.540*** (0.572)	7.034*** (0.402)	4.290*** (0.260)	4.290*** (0.260)
Level-2	Final overall symptoms (BDI-II)	σ_0^2	14.213*** (3.403)	18.806*** (3.390)	38.315*** (6.956)	23.423*** (4.526)	27.749*** (5.188)
	<i>Linear term</i>						
	Variance	σ_1^2		2.716*** (0.603)	2.106*** (0.455)	1.711* (0.430)	2.070*** (0.453)
	Covariance with overall symptoms (BDI-II)	σ_{01}^2		0.319*** (0.065)	0.353 (1.098)	-3.374** (2.286)	0.280~ (1.121)
	<i>Quadratic term</i>						
	Variance	σ_2^2		0.020*** (0.004)	0.133*** (0.048)	0.020*** (0.004)	0.020*** (0.004)

Covariance with overall symptoms (BDI-II)	σ_{02}^2	-0.232* (0.107)	-1.058*** (0.336)	-0.240* (0.111)	-0.227* (0.103)
<i>Cubic term</i>					
Variance	σ_3^2		0.001** (0.017)		
Covariance with overall symptoms (BDI-II)	σ_{03}^2		-0.053* (0.022)		
Pseudo R ²		0.827	0.885	0.930	-
Statistics					
Goodness-of-fit	Deviance	5284.750	4090.239	3924.481	3685.998
	AIC	5290.750	4098.239	3936.481	3705.998
	BIC	5304.603	4116.709	3964.186	3752.172
					3685.619
					3709.619
					3765.028
					3766.299

~ $p < .10$; * $p < .05$; ** $p < .01$; *** $p < .001$ significance level (2 tailed Pearson correlation).

In comparison, when accounting for time in the unconditional fixed linear growth model (Model 2), the estimated mean intercept value at Session 12 (γ_{00}) is estimated to be 11.05 ($p < .001$) for BDI-II predictions. Model 2 also introduces the γ_{10} parameter which estimates the fixed estimate of the rate of change. For BDI-II scores, Model 2 estimates a true rate of change of -2.143 for BDI-II scores. Model 2 thus indicates that on average, clinicians predict Mr T's BDI-II score to decrease by 2.143 points per session. In addition to these fixed parameter estimates, Model 2 provided an estimate of how much variance the effect of sessions accounts for. As is evident in Table 6.1, a Level-1 drop in σ_{ϵ}^2 between Model 1 and Model 2 indicates that clinicians have variance in their BDI-II scores and a large portion of that variance can be explained by Session effects, with the Pseudo R^2 (R^2_{ϵ}) indicating that 84% of the error variance in Model 1 for BDI-II scores can be accounted for by Session.

The Level-2 variance, demonstrated in Model 2, is estimated by the σ_0^2 values. This value indicates how much the intercept and rate of change of overall symptom change varies between individuals. Table 6.1 shows that these values for Model 2 are significant ($p < .001$). Such results indicate that individuals are likely to have different intercepts and slopes, and thus future models should account for this by including random effects.

Whether Model 2 is a better fit to the data than Model 1 can be determined using the deviance and BIC statistics. As can be seen in Table 6.1, Model 2 for the BDI-II demonstrates a reduction in the deviance statistic of 1360.3 with this change being significant. The BIC further demonstrated a large reduction of 1187.9. As such, Model 2 was deemed as a significantly better fit than Model 1.

Model 3 introduced a random linear effect of Session and also allowed the model to account for the data's covariance structure. As indicated by Singer and Willett (2003) and Field (2009) model fit can be improved by taking into account the appropriate covariance structure. Furthermore, this option has been used in a number of longitudinal studies and can often offer the best fit (Shek & Ma, 2011). While the AR(1) covariance structure has been identified as the most appropriate covariance structure to apply to repeated measures data, using this covariance structure on the BDI-II data resulted in no improvement in goodness-of-fit statistics. While the unstructured model requires the most parameters, which can result in lower statistical power, this study's purpose of identifying a substantively better fitting model means that this issue is not as predominant. As such, and based on recommendations from Heck and colleagues (2010), growth models were completed using the unstructured covariance structure as was used for Model 1 and Model 2. Like Model 2, Model 3 includes the linear effect of Session and, as such, the fixed estimates are the same across both models.

The addition of random effects to the linear growth model again provides the σ_{ϵ}^2 and σ_0^2 estimates, adding estimates for the between-person residual variability for the predicted rate of change (σ_1^2) and the covariance between the true intercept and true slope (σ_{01}^2). In Table 6.1, the Level-2 variance components for Model 2, σ_0^2 and σ_1^2 , were significant ($p < .001$), meaning that clinicians have different intercepts and slopes. As such, future models should account for this by including random effects. The final variance component value, σ_{01}^2 , addresses the population covariance between the true intercept (σ_0^2) and true rate of change (σ_1^2). Model 2 provides a significant positive covariance statistic for BDI-II predictions indicating that clinicians who predict a more improved outcome at sessions 12 on the BDI-II, also predict a quicker rate of progress

over the course of therapy. The inclusion of random effects demonstrates an improved model fit based on statistically significant change in deviance and BIC values, further supporting the notion that accounting for differences between clinicians in predictions of final score and rate of change provides a better fit to the data.

As Model 2 and 3 were based on linear growth, Model 4 incorporated the effect of Session² to test if adding a quadratic polynomial significantly improved model fit. Table 6.1 shows the quadratic (γ_{20}) effect for Model 4 is significant ($\gamma_{20} = 0.09$; $p < .001$), showing that BDI-II predictions are better accounted for by a quadratic trend. This value further indicates that the true rate of change for Model 4 is 0.09. This trend is further supported by the deviance and BIC statistic, with the deviance statistic for Model 4 indicating significant ($p < .05$) reductions for BDI-II predictions using the appropriate critical chi-square statistic. Re-running the analysis with the rescaled variable for time (i.e. reversed), Model 4 indicates that the instantaneous rate of change in the predictions of overall symptom change (i.e. BDI-II scores), as denoted by γ_{10} , was -1.20 at the end of therapy, beginning at -3.27 at session 0. Using the original and recoded γ_{00} , the average level of overall symptom change at session 12 is predicted to be 12.5.

In terms of the variance components, and as evidenced in Table 6.1, the Level-1 drop in σ_{ϵ}^2 between Model 3 and Model 4 alongside the Pseudo R² (R^2_{ϵ}) shows that taking into account quadratic effect of time explains 39% of the variance for predictions of overall symptoms. The remaining σ_0^2 , σ_1^2 , and σ_2^2 variance components for Model 4 estimate whether clinicians vary in their quadratic predictions of time. Table 6.1 shows that all three of these values are positive, non-zero and significant ($p < .001$). In addition, the significant and positive covariance between the true intercept and true quadratic slope across individuals, denoted by σ_{01}^2 , indicates that, like Model 3, clinicians in Model 4

who predict a higher BDI-II score at session 12 predict a slower rate of progress over the course of therapy.

Incorporating the effect of Session³ (i.e. cubic trend) as a fixed and random effect, Model 5 tested whether a cubic polynomial significantly improved the fit of the regression model. As can be seen in Table 6.1, γ_{30} demonstrates a significant effect ($p < .05$) and significant deviance and BIC statistics, indicating that BDI-II can be modelled as a cubic progression. However, the magnitude of the variance explained by cubic predictions is extremely small, thus suggesting that a quadratic growth model explains BDI-II predictions better given that a more parsimonious model is preferable.

Having identified the most appropriate trend for the data, Model 6 introduced the effects of experience level as a Level-2 predictor for the linear effect of time. As Model 4 was deemed as the best fitting growth model, Model 6 parameters in Table 6.1 are compared to Model 4 as opposed to Model 5. With trainee clinicians coded as 0 and experienced clinicians coded as 1, the γ_{00} value estimates the BDI-II score at session 12 for trainee clinicians (12.65, $p < .001$), with the γ_{01} value estimating the differential in the estimated BDI-II score at session 12 for between trainee and experienced clinicians. As shown in Table 6.1, the difference, -0.37, was not statistically different from zero. The estimated rate of change in BDI-II for trainee clinicians was denoted by γ_{10} (-1.160, $p < .001$). The Model 6 difference between the rates of change predicted by trainee and experienced clinicians, denoted by γ_{11} , indicated a value indistinguishable from zero (-0.08 *ns*). The non-significance of the values including the effect of experience shows that there is little difference between experience level and BDI-II predictions. This is further confirmed by a improvement of model fit that was not significant represented by the deviance statistic (0.68, *ns*) and the BIC (-12.86, *ns*). This indicates that when taking into

account experience on the linear effect of change, clinicians do not differ significantly between their BDI-II predictions over sessions based on their experience level.

Model 7 introduced the effects of experience level as a Level-2 predictor for the quadratic effect of time with γ_{00} estimating the BDI-II score at session 12 for trainee clinicians to be 12.36 ($p < .001$). The difference between trainee and experienced clinicians, denoted by γ_{01} was 0.20, and indistinguishable from zero. The rate of change at the end of therapy for trainees, denoted by γ_{10} (-0.75 $p < .01$), demonstrated differences between clinician experience level (-0.90, $p < .05$). The average rate of quadratic change in BDI-II predictions for trainee clinicians in Model 7 was deemed significant (0.14, $p < .001$), with a significant difference for experience levels (-0.04, $p < .05$). The impact of experience on predictions of change were overall confirmed as not significant ($p < .05$) when considering goodness-of-fit statistics when comparing Model 7 to Model 4 (the best fitting model in the analysis thus far), with the deviance statistic demonstrating a non-significant decrease, and the BIC values increasing between each model, showing reduced fit as opposed to improvement. As with Model 6, and as can be seen in Table 6.1, the effect of experience on the quadratic growth model demonstrated similar results for PANAS-NA and number of activities predictions, indicating that there is no significant difference in the shape of change for BDI-II, PANAS-NA and behavioural predictions between trainee clinicians and experienced clinicians.

Given that there was significant inter-individual variance in the MLM analyses, this finding was further explored using simple ordinary least squares (OLS) regression. OLS identified the number of experienced and trainee clinicians that predicted a linear, quadratic, or cubic progression of change on the BDI-II scores across therapy. Results indicated that, based on significant R^2 values, the shape of change was best described as a

linear progression for 34 (50%) of clinicians, a quadratic trend for 28 (41%) of clinicians and a cubic trend for 6 (9%) of clinicians.

Model B: Negative affect

The parameters for each model developed in the multilevel modelling of the predictions of PANAS-NA scores are provided in Table 6.2 (pp. 132-133).

For predictions using the PANAS-NA scores at session 2, Model 1 estimates the average intercept (i.e. if we averaged all scores across all sessions), denoted by γ_{00} , to be 17.24 ($p < .001$).

Variance components at both Level 1 (within-person) and Level 2 (between-person) are addressed by σ_{ε}^2 and σ_0^2 respectively. In Model 1 of the MLM analysis for negative affect scores, these were both significant at the $p < .001$ level. PANAS-NA predictions demonstrated an ICC of 0.226, indicating between individual differences accounting for 23% of all variability. As such, this analysis indicates that it is worthwhile to explore the variation in predictions made with the PANAS-NA between level of experience.

Table 6.2.
Results of Fitting a Taxonomy of Multilevel Models for Change on the PANAS-NA Across the Course of Therapy

Parameter	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7
Fixed effects							
Initial status, π_{0i}	γ_{00}	17.238*** (0.401)	10.318*** (0.431)	10.318*** (0.580)	11.248*** (0.532)	11.406*** (0.493)	11.043*** (0.753)
	γ_{01}					0.411 (1.064)	0.398 (1.064)
Rate of change, π_{10}	γ_{10}	-1.384*** (0.032)	-1.384*** (0.067)	-0.764*** (0.166)	-0.513** (0.188)	-0.735*** (0.179)	-0.587* (0.233)
	γ_{20}			0.062*** (0.016)	0.128* (0.055)	0.062*** (0.016)	0.077*** (0.023)
	γ_{30}				0.004 (0.004)		
	γ_{11}					-0.057 (0.134)	-0.354 (0.330)
	γ_{12}						-0.030 (0.031)
Variance components							
Level-1	Within-person	28.517*** (1.547)	7.448*** (0.404)	4.546*** (0.260)	2.949*** (0.180)	2.390*** (0.155)	2.949*** (0.180)
Level-2	Initial negative affect (PANAS-NA)	8.317*** (1.876)	10.232*** (3.919)	21.400*** (3.919)	17.560*** (3.894)	14.670*** (2.842)	17.520*** (3.300)
	<i>Linear term</i>						
	Variance			0.264*** (0.052)	1.513*** (0.324)	0.851* (0.427)	1.503*** (0.323)
	Covariance negative (PANAS-NA)			1.750*** (0.400)	1.299 (0.770)	-1.362 (0.772)	1.336~ (0.772)
	with affect (PANAS-NA)						-1.335~ (0.763)
	<i>Quadratic term</i>						

Variance	σ_2^2	0.013*** (0.003)	-0.587** (0.227)	-0.045 (0.069)	-0.045 (0.068)
Covariance negative (PANAS-NA) <i>Cubic term</i> Variance	σ_{02}^2	-0.048 (0.068)	0.114*** (0.035)	0.013*** (0.003)	0.013*** (0.003)
Covariance negative affect (PANAS-NA)	σ_{03}^2		0.032~ (0.017)		
			-0.001*** (0.000)		
Pseudo Statistics	R²				
		0.390	0.351	-	-
Goodness-of-fit	Deviance	4726.624	3610.051	3349.456	3425.557
	AIC	4732.624	3622.051	3379.456	3449.557
	BIC	4746.477	3649.755	3448.717	3504.966
	R^2_ϵ	0.739	0.390	-	-

~ $p < .10$; * $p < .05$; ** $p < .01$; *** $p < .001$ significance level (2 tailed Pearson correlation).

In contrast, Model 2 accounted for time and random intercepts (i.e. the unconditional fixed linear growth model), with the mean intercept value at Session 12 (γ_{00}) estimated to be 10.32 ($p < .001$) for PANAS-NA predictions. The fixed estimate of the rate of change, denoted by γ_{10} , was estimated to be -1.38 ($p < .001$). Model 2 thus indicates that on average, clinicians predict Mr T's PANAS-NA scores to decrease by 1.38 points per session. When considering the amount of variance that Session accounts for, the Level-1 drop in σ_{ϵ}^2 between Model 1 and Model 2 in Table 6.2 indicates that clinicians have variance in their PANAS-NA scores and a large portion of that variance can be explained by Session effects, with the Pseudo R^2 (R^2_{ϵ}) for PANAS-NA scores indicating that 89% of variance can be accounted for by Session.

How much the intercept and rate of change vary when considering change in negative affect scores is estimated by the σ_0^2 value. As can be seen in Table 6.2 this value for Model 2 is significant ($p < .001$) indicating that there is variation in how scores change over sessions across clinicians. The deviance statistic for Model 2 developed for PANAS-NA scores (Table 6.2) provided a reduction between Model 1 and Model 2 of 1116.6, deemed significant when compared to the appropriate critical chi-square statistic, indicating a significant improved model fit. Furthermore, the BIC demonstrated a reduction of 1096.7 between Model 1 and 2.

Model 3 introduced a random linear effect of Session and allowed the model to account for the data's covariance structure. Again, the unstructured covariance structure was applied. Due to Model 3 including the same fixed effects as Model 2, the fixed estimates remain the same across both models. The Level-1 and Level-2 variance components for Model 2 of the analysis of PANAS-NA scores (Table 6.2)

were identified as significant ($p < .001$) meaning that, as with overall symptoms, clinicians have different intercepts and slopes in terms of their predictions of negative affect change across 12 sessions of CBT. As such, future models should account for this by including random effects. The population covariance between the true intercept (σ_0^2) and true rate of change (σ_1^2) denoted by, σ_{01}^2 , provides a significant positive covariance statistic for PANAS-NA predictions indicating that, put simply, clinicians who predict a more improved outcome at sessions 12, also predict a quicker rate of progress over the course of therapy. The inclusion of random effects demonstrates an improved model fit based on statistically significant change in deviance and a substantial change in BIC values, further supporting the notion that accounting for differences between clinicians in predictions of final score and rate of change provides a better fit to the data.

Model 4 modelled the effect of Session² to test if a quadratic polynomial significantly affected the models of PANAS-NA scores. Table 6.2 indicates that the quadratic (γ_{20}) effect for Model 4 is significant ($\gamma_{20} = 0.06; p < .001$), showing that a quadratic trend fits PANAS-NA data better than a linear trend. The overall true rate of change for the quadratic growth model for PANAS-NA scores is estimated to be 0.06. Re-running the analysis with sessions reversed indicated that the instantaneous rate of change (γ_{10}) in the predictions of negative affect decreases in magnitude from -2.13 at session 0 to -0.51 at the end of therapy. Based on the original and recoded γ_{00} , the average level of overall symptom change at session 2 is predicted to be 27.15, dropping to 11.25 by session 12. The variance components for Model 4 indicate that the quadratic effect of time explains 35.1% of the variance for predictions of negative affect. To analyse whether clinicians vary in their predictions

of intercepts and rate of change, the values of σ_0^2 , σ_1^2 , and σ_2^2 were reviewed, outlined in Table 6.2. These values all demonstrated significant values. The covariance statistic for Model 4, denoted by σ_{01}^2 , indicates that clinicians who predict a higher PANAS-NA score at the end of therapy also predict a slower rate of change over the course of therapy. The deviance and BIC values Model 3 and Model 4 to confirm a significant reduction (deviance; $p < .05$) or substantial change (BIC; difference > 10), indicating that the random effects quadratic growth model provided a better fit than a random effects linear growth model.

As can be seen in Table 6.2, the effect of the cubic modelling of time (Model 5) demonstrates a non-significant effect for PANAS-NA scores. As Model 4 provided a better fit to the negative affect predictions, this model will be used for further analysis.

When considering the effect of experience level on linear prediction (Model 6), the non-significance of the values for γ_{01} (0.41, *ns*) and γ_{11} (-0.06, *ns*), as denoted in Table 6.2, indicated that experience level has little to no effect on the variability of predictions on linear growth. The deviance statistic and BIC also demonstrated non-significant/non-substantial change between Model 4 and Model 6, indicating that modelling these effects did not provide a better fit to the data.

Applying the effect of experience to both linear and quadratic predictions (Model 7) again yielded non-significant values for γ_{01} (0.40, *ns*) and γ_{11} (-0.35, *ns*), and the additional γ_{12} (-0.03, *ns*) indicating a lack of effect for experience on PANAS-NA predictions.

As a supplement to the remaining unexplained variance, simple linear regression was used to identify the frequency of linear, quadratic, and cubic predictions of PANAS-NA scores across 12 sessions of therapy for each participant. For predictions of negative affect 34% (23) of all clinicians predicted a linear trend, 40% (27) predicted quadratic trend, 25% (17) predicted a cubic trend, with one clinician's prediction demonstrated no significant trend.

Model C: Behavioural change

The parameters for each model developed in the multilevel modelling of predictions of the number of activities participated in between sessions are provided in Table 6.3 (pp. 138-139)

As demonstrated in Table 6.3, γ_{00} indicates that the average number of activities participated in across all sessions was estimated to be 4.179 ($p < .001$). In addition, the variance components at both Level 1 (σ_{ϵ}^2 ; within-person) and Level 2 (σ_0^2 ; between-person) were identified as significant at the $p < .001$ level. With a calculated ICC of 0.304, Model 1 indicated that between individual differences account for 30% of all variability across predictions of behavioural change. Again, these ICC values show that it may be beneficial to explore intra-individual differences.

Table 6.3.
Results of Fitting a Taxonomy of Multilevel Models for Change in the Number of Activities Participated across the Course of Therapy

Parameter	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7
Fixed effects							
Initial status, π_{0i}							
Intercept	4.179*** (0.182)	6.830*** (0.192)	7.830*** (0.263)	6.403*** (0.251)	6.442*** (0.623)	6.720*** (0.351)	6.674*** (0.352)
Exp						-0.635 (0.496)	-0.543 (0.480)
Rate of change, π_{10}							
Session		0.530*** (0.012)	0.530*** (0.026)	0.246*** (0.066)	0.308** (0.104)	-0.273*** (0.071)	0.140 (0.092)
Session ²				-0.028*** (0.006)	-0.012 (0.023)	-0.028*** (0.006)	-0.042*** (0.008)
Session ³					0.001 (0.001)		
Exp: Session						-0.054 (0.051)	0.212* (0.130)
Exp: Session ²							0.026* (0.026)
Variance components							
Level-1							
Within-person	σ_ϵ^2	4.259*** (0.231)	1.165*** (0.063)	0.752*** (0.043)	0.500*** (0.030)	0.441*** (0.029)	0.500*** (0.030)
Level-2							
Final # of activities participated in	σ_0^2	1.860*** (0.386)	2.141*** (0.385)	4.452*** (0.805)	3.994*** (0.735)	4.232*** (0.786)	3.920*** (0.722)
<i>Linear term</i>							
Variance	σ_1^2			0.038*** (0.008)	0.234*** (0.051)	0.448*** (0.128)	0.222*** (0.049)
Covariance with # of activities participated in	σ_{01}^2			0.321*** (0.070)	0.298* (0.146)	0.538* (0.243)	0.327* (0.143)
<i>Quadratic term</i>							

Variance	σ_2^2	0.002*** (0.000)	0.019** (0.006)	0.002*** (0.004)	0.002*** (0.004)
Covariance with # of activities participated in	σ_{02}^2	-0.004 (0.107)	0.059 (0.050)	-0.001 (0.013)	-0.001 (0.012)
<i>Cubic term</i> Variance	σ_3^2		0.000** (0.000)		
Covariance with # of activities participated in	σ_{03}^2		0.004 (0.003)		
Pseudo R ²	R ²	0.726	0.355-	-	-
Goodness-of-fit	Deviance	3326.251	2261.199	2084.776	2060.486
	AIC	3332.251	2273.199	2104.776	2090.486
	BIC	3346.104	2471.407	2150.950	2159.747
				2083.267	2107.267
				2162.675	2164.465

~ $p < .10$; * $p < .05$; ** $p < .01$; *** $p < .001$ significance level (2 tailed Pearson correlation).

When accounting for time and random intercepts in the unconditional fixed linear growth model (Model 2), the estimated mean intercept value at Session 12 (γ_{00}) is estimated to be 6.83 ($p < .001$) for the predicted number of activities Mr T participated in. As identified in Table 6.3, the fixed estimated rate of change (γ_{10}) was identified as 0.52. Model 2 thus indicates that on average, number of activities participated in increase by 0.53 at each session. In Table 6.3, the Level-1 drop in σ_{ϵ}^2 between Model 1 and Model 2 indicates that clinicians have variance in their predictions of the number of activities Mr T participated in and a large portion of that variance can be explained by Session effects, with the Pseudo R^2 (R^2_{ϵ}) indicating that 82% of the error variance in Model 1 for BDI-II scores can be accounted for by Session. The reduction in the deviance statistic between Model 1 and Model 2 for number of activities participated in was 1065.1 with a reduction in BIC of 1045.2. Reductions in the deviance statistic between Model 1 and Model 2 were significant at the $p < .01$ level as judged by the critical chi-square statistic, with the BIC difference being greater than 10, highlighting a substantial change.

Model 3 introduced a random linear effect of Session and allowed for the data's covariance structure to be accounted for. Again, the AR(1) covariance structure was applied and resulted in a non-zero change in goodness-of-fit statistics. As such, Heck and colleagues' (2010) recommendations were followed and the unstructured covariance structure used for the proceeding growth models. Like Model 2, Model 3 includes the linear effect of Session and, as such, the fixed estimates are the same across both models. The addition of random effects to the linear growth model again provided the σ_{ϵ}^2 and σ_0^2 estimates. As can be seen in Table 6.3 both of these values indicated that

clinicians demonstrate significant variability in their predictions and that predictions of the final session score differ between clinicians. In addition, the estimates for the between-person residual variability for the predicted rate of change (σ_1^2) was identified as significant ($p < .001$), indicating that clinicians are likely to predict different rates of change in their predictions of number of activities Mr T participated in. Furthermore, future models should account for this by including random effects. For Model 2, the covariance between the true intercept and true rate of change (σ_1^2) yielded a significant positive outcome indicating that clinicians, who predict a higher number of activities participated in at sessions 12, also predict a quicker rate of progress over the course of therapy. The inclusion of random effects demonstrates an improved model fit based on statistically significant change in deviance (176.4, $p < 0.01$) and a substantial change in BIC (150.0, $p < 0.01$) values, further supporting the notion that accounting for differences between clinicians in predictions of final score and rate of change provides a better fit to the data.

To understand whether a quadratic polynomial model provided a significant better fit than a linear model, Model 4 modelled the effect of Session². The quadratic effect, as denoted by γ_{20} in Table 6.3 demonstrated a significant effect ($p < .001$), indicating that a quadratic trend fits predictions of behavioural change better than a linear trend. The overall true rate of change for the quadratic growth model of the number of enjoyable behaviours participated in is estimated to be -0.03. Using the γ_{10} identified when re-running the analysis with reversed sessions, the instantaneous rate of change for behaviour was estimated to decrease in magnitude from 0.87 to 0.31. Based on the original and recoded γ_{00} , behavioural change was expected to improve from 0.26

at session 2 to 6.40 at session 12. Model 4's variance components shows that the quadratic effect of time accounts for 34% of the variance for predictions of number of activities Mr T participated in. The variance component values of σ_0^2 , σ_1^2 , and σ_2^2 in Table 6.3 indicate that clinician predictions of intercept and rate of change are likely to vary within and between each other. Model 4 further indicates, using σ_{01}^2 , that those who predict a lower level of improvement at session 12, predict a slower rate of change. Superior model fit is achieved by assessing the deviance for significant reductions ($p < .05$) using the appropriate chi-square critical statistic and BIC statistics for a substantial change (difference > 10). Results indicated that a quadratic growth model (Model 4) provided a superior fit to a linear growth model (Model 3).

Incorporating the effect of Session³ as a fixed and random effect, Model 5 tested whether a cubic polynomial significantly improved the fit of the regression model for the number of activities Mr T participated in. As can be seen in Table 6.3, γ_{30} demonstrates a non-significant effect ($p < .05$). This shows that clinicians predictions of behavioural change is not best represented using a cubic trend and, as such, reverting to Model 4 is likely to provide a better fit for future analyses.

Having established the best fitting growth model the effect of experience was applied to the data. Model 6 shows that clinician predictions of the number of activities Mr T participated in are not dependant (i.e. not significant) on experience level as determined by the values for γ_{01} (-0.63, *ns*) and γ_{11} (-0.05, *ns*). Furthermore the deviance statistic demonstrated non-significant change for predictions of behavioural

change, with the BIC yielding a non-substantial result, indicating that this model did not provide a better fit to the data.

Applying the effect of experience to both linear and quadratic predictions (i.e. Model 7) indicated that, based on the γ_{01} value (-0.54, *ns*) in Table 6.3, experience did not appear to have an effect on predictions on scores at sessions 12. The significance of γ_{11} (0.21, $p < .05$), and the additional γ_{12} (0.03, $p < .05$) indicated that there is an effect on predictions of the rate of change between clinicians. While changes in the deviance and statistic indicate a significantly improved fit (6.34, $p < 0.01$), it was outlined that the BIC was the preferred statistic to use in this analysis due to its conservative approach. The BIC (-13.52, difference > 10) indicates a non-substantial change and, as such, this model did not provide a superior fit.

As a supplement to the remaining unexplained variance, simple linear regression was used to identify the frequency of linear, quadratic, and cubic predictions of Mr T's participation in activities for each participant. Behavioural change predictions across therapy demonstrated a linear progression by 40 (59%), a quadratic trend by 21 (31%), and cubic by 5 (7%). Two clinicians provided predictions that did not significantly demonstrate either of the three possible trends.

Analysis 3: Presence of Discontinuous Patterns of Change

Using the definition of early rapid response provided by Ilardi and Craighead (1994), 9 (13%) clinicians predicted that Mr T would experience 70% of his overall change by session 5.

Based on the Tang and DeRubeis' (1999b) modified criteria, no clinicians predicted a sudden gain. While 28 of all participants predicted score changes between two sessions on the BDI-II that met the criteria for a 7 point drop between sessions that was 25% of the client's overall change, the third criterion, where the relative change is 1.5 times the individual's standard deviation of scores across therapy, was not met by any of the predictions. Three (4%) clinicians predicted that Mr T would experience a depression spike as defined by Hayes and colleagues (2007a).

Analysis 4: Continuous Change versus Discontinuous Change

When visually inspecting the curves produced from clinician predictions a number of patterns were identified that reflected smooth, continuous change, or variable, discontinuous change across the course of therapy. Smooth and continuous change was identified as change that did not demonstrate any noticeable deviations from a visibly smooth linear or quadratic trend. Variable, discontinuous change was identified as any change that showed at least two noticeable deviations from a smooth progression (i.e. a trend that has two inflection points). As such, cubic trends were automatically identified as a variable based on their definition meaning that the trend of change goes in two directions. Based on the above definitions 44 (65%) clinicians predicted overall symptom change, based on BDI-II scores, to be smooth and continuous, with 24 (35%) predicting variable, discontinuous change. A chi-square goodness of fit test showed a significant difference in frequencies of each change type, $\chi^2 (1) = 5.88$, $p = < 0.05$. Change in negative affect (PANAS-NA scores) was predicted to be smooth and continuous by 36 (53%) clinicians and predicted to be variable and discontinuous by 32

(47%) thought this difference was not significant $\chi^2 (1) = 0.24$ $p = .628$. Behavioural change, as predicted by number of activities participated in between sessions, was expected to be smooth and continuous by 33 (49%) clinicians, with 35 (52%) predicting variable, discontinuous change. This difference was not significant by a chi-square goodness of fit test, $\chi^2 (1) = 0.059$, $p = .808$.

Analysis 5: The Relationship between Overall Symptom, Mood and Behavioural Change.

The number of clinicians who predicted each type of pattern can be seen in Table 6.4, below.

Table 6.4.
Frequency of Each Sequence of Measured Dimension Expected for a Client with Depression

Order of change	Frequency
Simultaneous improvement (SBM)	13
Overall symptom improvement then behavioural improvement then mood improvement (S-B-M)	1
Mood improvement then behavioural improvement then overall symptom improvement (M-B-S)	3
Behavioural improvement then overall symptom improvement then mood improvement (B-S-M)	1
Behavioural improvement then mood improvement then overall symptom improvement (B-M-S)	19
Mood then simultaneous overall symptom and behavioural improvement (M-SB)	1
Behavioural improvement then simultaneous overall symptom and mood improvement (B-SM)	8
Simultaneous overall symptom and behaviour improvement then mood improvement (SB-M)	5
Simultaneous mood and behavioural improvement then overall symptom improvement (MB-S)	17

Following on from understanding the nature of change clinicians predict (i.e. continuous vs. discontinuous), the type of change predicted to occur first for a depressed client was predicted to be behaviour by 28 (41%) clinicians. Symptom improvement was predicted to occur first by 1 (2%) clinician and 4 (6%) clinicians predicted a change in mood would occur before the other types of change. Simultaneous behaviour and overall symptom change was predicted by 5 (7%) clinicians and simultaneous behaviour and mood change was predicted by 17 (25%) clinicians.

Analysis 6: Clinically Significant Symptom Reduction by the Final Session of Therapy

When making predictions using the BDI-II, all clinicians predicted the average score at session 12 to be 12.63 (SD = 5.25). Using Jacobson and Truax's clinical significance classification method, 1 (2%) clinician predicted that Mr T would remain Unchanged, 2 (2.9%) predicted that he would have improved, and 65 (96%) expected Mr T to fall within the Recovered range. When reviewing predicted PANAS scores at session 12, the average score expected was 11.26 (SD = 4.13). When understanding the level of improvement at the final therapy session, 3 (4%) of the participants predicted Mr T to demonstrate no change. Analyses further indicated that 60 (88%) of clinicians made predictions of Mr T's outcome falling within the Improved range, and 5 (8%) predicting that Mr T's scores would reflect the Recovered classification. None of the predictions made on either of these two measures indicated that Mr T would experience no change in terms of the clinical significance classification.

In terms of behavioural change for Mr T, the average of the predicted number of activities Mr T participated in between session 11 and session 12 was 6.47 (SD = 2.13) with 21 (31%) clinicians predicting that Mr T would participate in 5 or less activities, and 47 (69%) participating in 6 or more activities.

Analysis 7: Maintenance of Therapeutic Change at 3- and 6-Month Follow-up

Using a repeated-measures ANOVA design, significant differences in change between session 12, 3 month follow-up and 6 month follow-up were tested. For predictions of overall depression symptoms (BDI-II), negative affect (PANAS-NA), and number of activities, Mauchly's test indicated that the assumption of sphericity was violated ($p < .05$), as can be seen in Table 6.5, below.

Table 6.5.

Significance of Mauchly's Test of Sphericity for Session 12, 3 Month Follow-up and 6 Month Follow-up for Overall Symptom Change (BDI-II), Negative Affect (PANAS-NA) and Number of Activities

Measure	Mauchly's W	Degrees of freedom	Significance
PANAS-NA	0.909	2	.045
BDI-II	0.795	2	.001
Number of Activities	0.696	2	.000

As the PANAS-NA and BDI-II demonstrated a Mauchly's test statistic over 0.75, the degrees of freedom were corrected using Huyn-Feldt estimates of sphericity. As the number of activities demonstrated a Mauchly's test statistic less than 0.75, the Greenhouse-Geisser correction was used.

The average of predicted scores made on the BDI-II, PANAS-NA, and number of activities for Session 12, 3 month follow-up and 6 month follow-up are summarised in Figure 6-3, below. Results of the repeated measures ANOVA indicated that predictions in overall symptom change at session 12, 3 month follow-up and 6 month follow-up demonstrated significant differences, $F(1.72, 113.70) = 3.37, p = .05$, with predictions in negative affect also demonstrating significant differences, $F(1.91, 126.16) = 3.12, p = .05$.



Figure 6.3. Average BDI-II, PANAS-NA and No. of Activities score for Session 12, 3 month follow-up and 6 month follow-up.

While Tukey's test is typically the post-hoc test applied to repeated measures when comparing pairs of means, when sphericity has been violated, it is recommended that Bonferroni's adjustment be applied (Field, 2009). As such, pairwise comparisons were determined using the Bonferroni adjustment, with no significant comparisons

apparent. This is likely due to Bonferroni's test being more conservative than the ANOVA.

Change in the number of activities Mr T participated in over the follow-up sessions were shown to be not significant, $F(1.53, 101.23) = 0.64, p = 0.49$. The inclusion of experience to all of the above analyses as a between subject factor indicated that there was no interaction between Session and experience level.

To gather a comprehensive understanding of the type of change clinicians predict over the follow-up sessions, the frequency of deterioration and improvement, alongside the magnitude of such change was investigated. For expected change between session 12 and 3 month follow-up, 16 (24%) clinicians predicted no change in overall symptoms scores (BDI-II), with 32 (47%) clinicians predicting Mr T to improve on average by 2.81 points. Mr T's overall symptoms were predicted to deteriorate by 20 (29%) clinicians, with an average increase in scores of 2.45 between session 12 and 3 month follow-up.

Seventeen clinicians (25%) predicted no change in Mr T's negative affect scores. In terms of the clinicians who predicted that Mr T's negative affect scores would improve (19, 28%), the average improvement was 2.16 points, with the maximum amount of change predicted being 10 points. The average rate of change predicted by those who expected Mr T's negative affect to deteriorate (32, 47%) was 2.84, with a maximum of 8 points reduction in PANAS-NA scores. The change in number of activities that Mr T was effectively participating in between session 12 and 3 month follow-up was predicted to remain constant by 26 clinicians (38%). The number of

activities Mr T participated in was expected to drop by 21 clinicians (31%) at an average rate of -1.48, with 21 (31%) clinicians expecting improvement at an average rate of 1.33.

In terms of differences in change between 3 and 6 month follow-up, 22 (13%) clinicians predicted that Mr T would demonstrate no change in overall symptoms (BDI-II). On average, those who identified improved scores for Mr T (27; 43%) indicated an improvement of on average 1.93 with those who predicted deterioration (17; 25%) expecting an average of 2.35 reduction in BDI-II scores. No change was expected in Mr T's negative affect scores by 35 clinicians (52%), with 20 (29%) predicting improvement at an average of 2.4 points. Clinicians predicting deterioration (13; 19%) indicated an average reduction of 2.46.

Over the three months between the two follow-up sessions, Mr T was expected to show no difference in the number of activities he was participating in by 44 (65%) clinicians. Improvement was expected by 15 (22%) clinicians at an average of 1.27, with deterioration predicted by 9 (13%) clinicians with a mean reduction of 1 activity per session.

When looking at the overall change between the final session (session 12) and 6 month follow-up, Mr T was expected to show no difference in his BDI-II scores by 11 (16%) clinicians. At a rate of 2.85, 39 (57%) clinicians predicted Mr T's overall symptoms to improve between session 12 and 6-month follow up. Deterioration was expected by 18 (27%) clinicians, at an average rate of 3 points. Over the six months between the final and follow-up session, 11 (16%) clinicians predicted no change in Mr

T's negative affect, with 27 (40%) clinicians expecting improvement at an average rate of 1.78 points. At a rate of 2.73 points, 30 (47%) clinicians expected deterioration in Mr T's presenting affect. Mr T's participation in activities was expected to remain the same by 33 (49%) clinicians. Activity participation was expected to improve by 18 (27%) clinicians at an average rate of 3 activities over the six month period, with 17 (25%) clinicians expecting participation to reduce, on average, by 1.41 points.

Clinically significant changes were assessed over both the Session 12 to 3 month period, the 3 month to six month period and overall between session 12 and 6 month follow-up. Results indicated that between session 12 and 3 month follow-up, 66 (97%) clinicians predicted that Mr T would remain within the same clinical significant classification based on BDI-II predictions. Two clinicians predicted improvement from Improved to Recovered. In comparison, 12 and 3 month follow-up score predictions on the PANAS-NA indicated that 57 (84%) clinicians predicted no change in clinical significant classification, with one clinician predicting that Mr T would improve from the Unchanged classification to Improved. In addition, 6 (9%) clinicians predicted Mr T to deteriorate from Improved to Unchanged, and two clinicians predicted Mr T to deteriorate from Recovered to Improved.

Analysis of changes in scores predicted at 3 and 6 month follow-up using the BDI-II indicated that 66 (97%) clinicians expected Mr T's clinical significance classification to demonstrate no change, with two predicting a change from Recovered to Improved. PANAS-NA predictions for 3 and 6 month follow-up showed that 63 (93%) predicted no change to clinical significant classification, with 3 (4%) clinicians

expecting an improvement from Unchanged to Improved, one expecting deterioration from Improved to Unchanged, and one predicting deterioration from Recovered to Improved.

When reviewing the overall frequency of changes between clinical significance criteria between session 12 and 6 month follow-up, BDI-II scores were predicted to demonstrate no change by 64 (94%) clinicians. Two clinicians predicted that Mr T's scores would change from Improved to Recovered, with two clinicians additionally predicting scores moving from the Recovered range to the Improved range. Negative affect, as determined by PANAS-NA scores, were expected to remain unchanged between session 12 and 6 month follow-up by 56 (82%) clinicians. Two clinicians predicted Mr T's negative affect to shift from Unchanged to Improved over this time period, with 5 clinicians predicting a shift from Improved to Unchanged, 2 predicting a shift from Improved to Recovered, and 3 predicting a shift from Recovered to Improved.

Ms S

Analysis 2: Shape of Change over Session 2 -12

The analysis completed in terms of understanding the expected shape of change for Ms S will proceed in the same format as was completed for Mr T.

Model D: Overall anxiety symptom change

The parameters for each model developed in the multilevel modelling of the predictions of BAI scores/overall symptoms are presented in Table 6.6 (pp. 154 – 155).

As with the growth modelling completed for Mr T, Model 1 and 2 represented the unconditional means model and unconditional fixed linear growth model respectively. Model 1, the unconditional means model, provided the fixed estimate of the intercept, γ_{00} , reflecting the average BAI score across all participants at session 0, estimated to be 33.20 ($p < .001$). In addition to the fixed effects, the variance components at both Level 1 (within-person) and Level 2 (between-person) were addressed by σ_{ϵ}^2 and σ_0^2 respectively. In Model 1, both these estimates are significant at the $p < .001$ level. To provide rationale to develop an unconditional fixed linear growth model based on indications of between-group variance, the ICC was calculated using the equation $p = \frac{\sigma_0^2}{(\sigma_0^2 + \sigma_{\epsilon}^2)}$. Model 1 developed for BAI scores demonstrated an ICC of 0.222, indicating that 22% of all the variability in BAI predictions is between clinicians, suggesting that it may be worthwhile to explore intra-individual differences.

Model 2, the unconditional growth model, was developed to account for sessions and random intercepts. The estimated mean intercept value at Session 12 (γ_{00}) is estimated to be 20.96 ($p < .001$) for BAI predictions. The fixed estimate of the rate of change, γ_{10} , estimates a true rate of change of -2.45 for BAI scores. This indicates that, on average, Ms S's overall symptoms are expected to improve at a rate of 2.45 BAI points between each session.

Table 6.6.
Results of Fitting a Taxonomy of Multilevel Models for Change on the BAI Across the Course of Therapy

Parameter	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7
Fixed effects							
Initial status, γ_{00}	33.201*** (0.682)	20.960*** (0.736)	20.960*** (1.038)	22.076*** (0.952)	21.843*** (0.924)	22.749*** (1.342)	22.750*** (1.342)
Exp γ_{01}						-1.347 (1.898)	-1.348 (1.898)
Rate of Session change, π_{10}		-2.448*** (0.049)	-2.448*** (0.109)	-1.705*** (0.290)	-2.074*** (0.355)	-1.592*** (0.309)	-1.696*** (0.409)
Session ² γ_{20}				0.074** (0.027)	-0.023 (0.085)	0.074** (0.027)	0.064 (0.038)
Session ³ γ_{30}					-0.006 (0.006)		
Exp: Session γ_{11}						-0.225 (0.217)	-0.017 (0.579)
Exp: Session ² γ_{12}							0.021 (0.054)
Variance components							
Level-1							
Within-person σ_e^2	84.245*** (4.569)	18.304*** (0.993)	10.419*** (0.596)	5.731*** (0.347)	4.608*** (0.300)	5.731*** (0.347)	5.731*** (0.347)
Level-2							
Final level of overall symptoms (BAI) σ_0^2	23.997*** (5.445)	29.991*** (5.430)	69.996*** (12.574)	58.358*** (10.581)	54.467*** (9.968)	57.905*** (10.503)	57.905*** (10.503)
<i>Linear term</i>							
Variance σ_1^2			0.717*** (0.120)	4.981*** (0.979)	5.534*** (1.480)	4.992*** (0.981)	4.981*** (0.979)
Covariance overall symptoms (BAI) with symptoms σ_{01}^2			5.721*** (1.200)	4.724* (2.389)	1.481 (2.753)	4.718* (2.384)	4.718* (2.380)
<i>Quadratic term</i>							
Variance σ_2^2				0.044***	0.320**	0.044***	0.044***

Covariance overall (BAI)	with symptoms	σ_{02}^2	(0.009)	(0.085)	(0.009)	(0.009)
			-0.110	-0.707	-0.103	-0.103
			(0.214)	(0.651)	(0.213)	(0.213)
<i>Cubic term</i>						
Variance		σ_3^2		0.001***		
				(0.000)		
Covariance overall (BAI)	with symptoms	σ_{03}^2		-0.038		
				(0.043)		
Pseudo R ² Statistics and Goodness-of-fit			R ² _{ϵ}	0.431	0.450	
Deviance			5535.656	4497.549	4247.135	3985.363
AIC			5541.656	4505.549	4259.125	4005.363
BIC			5555.508	4524.018	4286.840	4051.537
			3989.696	4063.580	3984.171	3984.023
					4008.171	4010.023
					4063.580	4070.049

~ $p < .10$; * $p < .05$; ** $p < .01$; *** $p < .001$ significance level (2 tailed Pearson correlation).

Model 2 additionally provides an estimate of how much variance the effect of Session accounts for. As is evident in Table 6.6, a significant Level-1 drop in σ_{ε}^2 between Model 1 and Model 2 shows that clinicians demonstrate variance in their BAI scores and a large portion of that variance can be explained by Session effects. Furthermore, the Pseudo R^2 (R^2_{ε}) indicates that 78% of the error variance in Model 1 for BAI scores can be accounted for by Session. Model fit was determined based on the deviance and BIC statistics. The change in these statistics in Table 6.6 for Model 2 demonstrates a reduction in the deviance statistic of 1038.107 and a reduction in the BIC of 1031.49. Deviance reductions between Model 1 and Model 2 were significant at the $p < .01$ level as judged by the critical chi-square statistic, indicating a significant improved model fit. BIC reductions proved substantial based on the criteria of a substantial difference reflecting a change of 10 or more.

Model 3 introduced a random linear effect of Session and allowed the model to account for the data's covariance structure. The unstructured covariance structure was applied as informed by Heck and colleagues (2010). As Model 3 added the random effect of time, the fixed parameters remain the same as Model 2. The addition of random effects to the linear growth model includes the covariance parameters σ_{ε}^2 and σ_0^2 along with the between-person residual variability estimates for the predicated rate of change, σ_1^2 , and the covariance between the true intercept and true slope, σ_{01} . Both the Level-1 and Level-2 variance components for Model 3 were identified as positive, non-zero and significant ($p < .001$), as can be seen in Table 6.6, meaning that clinicians are likely to have different intercepts and slopes. As such, future models should account for this by including random effects. The final variance component value, σ_{01}^2 , addresses the population covariance between the true intercept (σ_0^2) and true rate of change (σ_1^2). Model 3 provides a significant positive

covariance statistic for overall symptoms change, indicating that clinicians who predict a higher score at sessions 12 on the BAI, also predicted a slower rate of progress over the course of therapy. The inclusion of random effects demonstrates an improved model fit based on statistically significant change in deviance (250.4, $p < 0.01$) and BIC (237.2, difference > 10) values, further supporting the notion that accounting for differences between clinicians in predictions of final score and rate of change provides a better fit to the data.

As Model 2 and 3 were based on linear growth, Model 4 incorporated the effect of Session^2 to test if a quadratic polynomial significantly affected the models of BAI scores. Table 6.6 shows the quadratic effect (γ_{20}) for Model 4 is significant ($p < .001$), indicating that BAI scores are better accounted for by a quadratic trend. The overall true rate of change for the quadratic growth model is estimated as 0.07 for the BAI model. Re-running the analysis with the reversed variable for time, Model 4 indicates that the instantaneous rate of change in the predictions of overall symptom change (i.e. BAI scores), as denoted by γ_{10} , decreases in magnitude from -3.34 at session 0 to -1.71 at the end of therapy. Using the original and recoded γ_{00} , the average level of overall symptom change at session 2 is predicted to be 49.83, dropping to 22.08 by session 12. In terms of the variance components, and as evidenced in Table 6.6, the Level-1 drop in σ_{ϵ}^2 between Model 3 and Model 4 alongside the Pseudo R^2 (R^2_{ϵ}) indicates that taking into account quadratic effect of time explains 45% of the variance for predicted BAI scores. The remaining variance components, σ_0^2 , σ_1^2 , σ_2^2 , for Model 4 estimate whether clinicians vary in their quadratic predictions of time. Table 6.6 shows that these values for overall symptoms are significant ($p < .001$). In addition, the significant covariance between the true intercept and true quadratic slope across individuals, denoted by σ_{01}^2 , indicates that, like Model 3, clinicians in Model 4 who predict a

higher BAI score at session 12 predict a slower rate of progress over the course of therapy. Whether the quadratic effect of Model 4 provides a better fit to the data than Model 3 can be understood using the deviance and BIC statistics. The deviance statistic for Model 4 indicates significant ($p < .05$) reductions for Ms S's overall symptoms using the appropriate critical chi-square statistic, with the BIC confirming a substantial change (i.e. > 10).

Incorporating the effect of Session³ (i.e. cubic trend) as a fixed and random effect, Model 5 tested whether a cubic polynomial significantly improved the fit of the regression model. As can be seen in Table 6.6 BAI score predictions demonstrated a non-significant effect of the cubic effect of time, thus indicating that cubic growth does not best represent the pattern of change of clinician predictions for all three measures. As such, the quadratic growth model (Model 4) demonstrates the best fit to the data in terms of the trend of predictions over time.

Following the identification of the most appropriate trend for the data, the effects of experience level as a Level-2 predictor for the linear effect of time was introduced to Model 6. As Model 4 was deemed as the best fitting growth model, Model 6 parameters are compared to Model 4 as opposed to Model 5. With trainee clinicians coded as 0 and experienced clinicians coded as 1, the γ_{00} value estimates the BAI score at session 12 for trainee clinicians (22.75, $p < .001$), with the γ_{01} value estimating the differential in the estimated BAI score at session 12 between trainee and experienced clinicians. As shown in Table 6.6, the difference, -1.35, was not significant. For Model 6, γ_{10} indicated the estimated rate of change in BAI for trainee clinicians, providing a result of -1.49 ($p < .001$). The Model 6 difference between the rates of change predicted by trainee and experienced clinicians,

denoted by γ_{11} , indicated a value indistinguishable from zero (-0.23, *ns*). The non-significant values of the effect of experience indicates that trainee and experienced clinicians demonstrate little difference in their predictions of BAI scores. This is further confirmed by improvement of model fit that was not significant represented by the deviance statistic (1.19, *ns*) and non-substantial, provided by the BIC (-12.04). This shows that when taking into account experience on the linear effect of change, clinicians do not differ significantly between their BAI predictions based on their experience level.

Model 7 introduced the effects of experience level as a Level-2 predictor for the quadratic effect of time. As can be seen by γ_{01} values, the difference between trainee and experienced clinicians was not significant and not distinguishable from zero in scores predicted at session 12. The rate of change toward the end of therapy for trainees, denoted by γ_{10} (-1.70, $P < .01$), also demonstrated non-significant differences between clinician experience level ($\gamma_{11} = -0.02$). The impact of experience on predictions of change were overall confirmed as not significant ($p < .05$) when considering goodness-of-fit statistics. When comparing the change in deviance (1.34, *ns*) and BIC (-18.51) statistics between Model 4 (the best fitting model in the analysis thus far) and Model 7, no significant or substantial change is apparent. In terms of the effect of experience on the quadratic growth model for PSWQ scores and percentage of activities completed, similar results were demonstrated as can be seen in Table 6.6. As such, the results of Model 7 indicate that there is no significant difference in BAI, PSWQ scores and behavioural predictions between trainee and experienced clinicians in predictions of change that reflect a quadratic growth pattern.

As a supplement to the remaining unexplained variance, simple linear regression was used to identify the frequency of linear, quadratic, and cubic predictions for predictions of overall symptoms in the hypothetical case of Ms S across the course of therapy for each participant. Results indicated that, based on significant R^2 values, predictions of overall symptom change (BAI) best reflected a linear progression by 34 (50%) clinicians, 22 (32%) predicted a quadratic trend, and 11 predicted a cubic trend, while one clinician's prediction demonstrated no significant trend.

Model E: Level of anxiety

The parameters for each model developed in the multilevel modelling of the predictions of PSWQ scores/level of anxiety are presented in Table 6.7 (pp. 162 – 163).

Model 1 provided an unconditional means model for predicted level of anxiety across the course of therapy. In terms of predicted PSWQ scores at session 2, Table 6.7 shows that Model 1 estimates the average intercept to be 52.10 (γ_{00} ; $p < .001$). Level-1 and Level-2 variance components are addressed in Table 6.7 with the values denoted by σ_{ϵ}^2 and σ_0^2 respectively. In Model 1, both these estimates are significant at the $p < .001$ level for predictions of anxiety level. The ICC was calculated using these values to identify whether it would be worthwhile exploring the effect of between-subject variance. The ICC for Model 1 developed for PSWQ predictions indicated between individual differences accounting for 23% of all variability. This ICC suggests that it may be beneficial to consider intra-individual differences.

Accounting for time and random intercepts, Model 2 provided an estimate of 34.78 (γ_{00} ; $p < .001$) for the average PSWQ score predicted at session 12. γ_{10} , the fixed estimate of

the rate of change, was identified as -3.47 ($p < .001$) indicating that overall, clinicians expected Ms S's level of anxiety to improve by 3.47 PSWQ points between each session. The Pseudo R^2 for PSWQ scores showed that 80% of the model's variance can be accounted for by Session. Model fit was determined based on the deviance and BIC statistics. The deviance statistic for Model 2 developed for PSWQ scores (Table 6.7) provided a reduction of 1095.38 and a reduction of 1088.76 for the BIC. As the deviance changes were significant at the $p < .01$ level when compared to the appropriate chi-square statistic, and the BIC change demonstrated a substantial result (i.e. > 10), Model 2 was deemed as a significantly better fit than Model 1.

Model 3 introduced a random linear effect of Session and allowed for the data's covariance structure to be accounted for. Again, the AR(1) covariance structure was applied and resulted in a non-zero change in goodness-of-fit statistics. As such, Heck and colleagues (2010) recommendations were followed and the unstructured covariance structure used for the preceding growth models. Fixed effects remained the same as Model 2 as random effects was the only addition to Model 3. The significance of the σ_ε^2 , σ_0^2 , and σ_1^2 in Table 6.7 indicate that clinicians are likely to predict different scores for sessions 12 and different rates of change in their expectations of change in anxiety levels. As such, this variability should be accounted for in future models. The significance of the covariance statistic (σ_{01}^2) indicates that a client who predicts a higher PSWQ score at session 12 is likely to predict a slower rate of change overall. Changes between Model 2 and Model 3's deviance indicate a significant change and BIC values indicate a substantial change, showing that the inclusion of random effects in a linear growth model provides a better fit than an unconditional linear growth model.

Covariance with level of anxiety (PSWQ)	σ_{02}^2	(0.013)	(0.192)	(0.013)	(0.013)
<i>Cubic term</i> Variance	σ_3^2	-0.657 (0.401)	-0.267 (1.537)	-0.736~ (0.399)	-0.728~ (0.393)
Covariance with level of anxiety (PSWQ)	σ_{03}^2		0.003*** (0.000)		
			0.028~ (0.102)		
Pseudo R ² Statistics and Goodness-of-fit	R ² _{ϵ}	0.800	0.417	0.364	
Deviance		6049.0200	4691.426	4472.960	4468.593
AIC		6055.020	4703.426	4492.960	4493.621
BIC		6068.872	4731.130	4539.134	4549.030
					4494.593
					4554.619

~ $p < .10$; * $p < .05$; ** $p < .01$; *** $p < .001$ significance level (2 tailed Pearson correlation).

Model 4 added the quadratic effect of time (Session²) to identify whether this model provided a better fit for the PSWQ data than a linear trend, with the significance ($p < .001$) of the γ_{20} value in Table 6.7 confirming this. The γ_{20} value further indicated that the overall true rate of change for the quadratic growth model for anxiety level across therapy is 0.12. Using the γ_{10} value, and re-running the analysis with sessions reversed, PSWQ scores were predicted to decrease in magnitude from -4.85 at session 0 to -2.31 at the end of therapy. Using both γ_{00} from the original and reversed session analyses, level of anxiety at session 0 was predicted to be 75.89 dropping to 36.51 by the end of therapy. The drop in values between Model 3 and Model 4's σ_{ε}^2 value, along with the (R^2_{ε}) value indicate that taking into account the quadratic effect of time explains 36% of the variance for predicted PSWQ scores. Model 4's remaining variance components, denoted by σ_0^2 , σ_1^2 , σ_2^2 , are identified in Table 6.7 as significant ($p < .001$), indicating that there is significant variation in intercept and rate of change predictions between clinicians. The significant covariance value (σ_{01}^2) shows that clinicians who predict higher scores at session 12 typically predict a slower rate of change. Model 4 was deemed to be a better fit than the random effects linear growth model as judged by a substantial change (difference between Model 3 and Model 4 BIC > 10) between the deviance and BIC statistics.

Model 5 introduced the modelling of the cubic effect of time (Session³). As can be seen in Table 6.7, the effect of this growth model produced a non-significant effect for PSWQ scores. As such, Model 4 provided a better fit to predictions of level of worry, and will thus be the growth model used for further analysis.

Having established the best fitting growth model the effect of experience was applied to the data in Model 6. Table 6.7 shows that clinician predictions of level of worry are not

dependent (i.e. not significant) on experience level as determined by the values for γ_{01} (-4.73, *ns*) and γ_{11} (-0.01, *ns*). The deviance statistic and BIC also demonstrated non-significant/non-substantial change for predictions of behavioural change indicating that this model did not provide a better fit to the data.

Applying the effect of experience to both linear and quadratic predictions again yielded non-significant values for γ_{01} (-4.26, *ns*) and γ_{11} (-1.00, *ns*), and the additional γ_{12} (-0.07, *ns*) indicating a lack of effect for experience on predictions of change in level of worry across a course of CBT.

As a supplement to the remaining unexplained variance, simple linear regression was used to identify the frequency of linear, quadratic, and cubic predictions of PSWQ scores across 12 sessions of therapy for each participant. Analyses indicated that 32 (47%) clinicians predicted Ms S's level of anxiety (PSWQ) to reflect a linear trend, with 20 (30%) clinicians predicting a quadratic trend and 16 (24%) clinicians predicting a cubic trend.

Model F: Percentage of Feared/Avoided Situations Completed

The parameters for each model developed in the multilevel modelling of the predicted percent of activities Ms S participated in/behavioural change are provided in Table 6.8 (pp. 166 – 167).

Table 6.8.
Results of Fitting a Taxonomy of Multilevel Models for Percent of Activities Participated In over the Course of Therapy

Parameter	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7
Fixed effects							
Initial status, π_{0i}							
Intercept	γ_{00}	41.902*** (2.119)	69.192*** (2.199)	69.192*** (3.029)	65.532*** (2.747)	65.726*** (2.680)	61.063*** (2.763)
Exp	γ_{01}					8.938~ (5.224)	7.638 (5.415)
Rate of change, π_{10}							
Session	γ_{10}	5.458*** (0.118)	5.458*** (0.273)	3.018*** (0.563)	3.326*** (0.827)	2.604*** (0.616)	2.162*** (0.782)
Session ²	γ_{20}			-0.244*** (0.052)	-0.163 (0.213)	-0.244** (0.052)	-0.291 (0.073)
Session ³	γ_{30}				0.005 (0.014)		
Exp: Session	γ_{11}					0.828 (0.533)	1.713 (1.106)
Exp: Session ²	γ_{12}						0.094 (0.103)
Variance components							
Level-1							
Within-person	σ_e^2	432.114*** (23.435)	104.440*** (5.664)	54.317*** (2.105)	35.052*** (2.125)	27.851*** (1.805)	35.052*** (2.125)
Level-2							
Final % of activities	σ_0^2	265.940*** (52.389)	294.728*** (52.348)	606.705*** (107.017)	492.758*** (88.005)	466.241*** (83.746)	478.595*** (85.644)
<i>Linear term</i>							
Variance	σ_1^2			4.557*** (0.867)	17.139*** (3.704)	28.210*** (8.055)	16.601*** (3.620)
Covariance with % of activities	σ_{01}^2			0.779*** (0.047)	12.189 (12.985)	-7.412 (18.318)	8.632 (12.622)
<i>Quadratic term</i>							
Variance	σ_2^2			0.143*** (0.032)	0.143*** (0.032)	2.033*** (0.533)	0.143*** (0.032)
Covariance with σ_{02}^2				-2.933	-2.933	-6.944	-3.113**

% of activities	(1.210)	(4.732)	(1.207)	(1.192)
<i>Cubic term</i>				
Variance		0.009*** (0.002)		
Covariance with % of activities		-0.259 (0.315)		
Pseudo R ² Statistics and Goodness-of-fit	R ² _ε	0.758	0.480	0.355
Deviance	6801.528	5835.875	5530.715	5325.889
AIC	6807.528	5843.875	5542.715	5345.889
BIC	6821.380	5862.344	5570.420	5392.063
			5259.068	5323.021
			5289.068	5347.021
			5358.329	5402.429
				5408.220

~ $p < .10$; * $p < .05$; ** $p < .01$; *** $p < .001$ significance level (2 tailed Pearson correlation)

Model 1 estimated the average percent of activities participated in to be 41.90 ($\gamma_{00}; p < .001$) as shown in Table 6.8. Level-1 and Level-2 variance components are addressed in Table 6.8 with the values denoted by σ_{ε}^2 and σ_0^2 respectively. In Model 1, both these estimates are significant at the $p < .001$ level. The ICC was calculated using these values to identify whether it would be worthwhile exploring the effect of between-subject variance with the between individual differences accounting for 30% of all variability in predictions of the percentage of activities Ms S participated in.

The unconditional growth model (Model 2) estimated the average of the percentage of activities Ms S participated in to be 69.192 ($p < .001$) at session 12, as denoted by γ_{00} in Table 6.8. In addition, Model 2 estimated that behavioural change increased by 5.46 on average ($\gamma_{10}; p < .01$) at each session. The Pseudo R^2 for and percentage of activities Ms S completed indicated that 76% of variance can be accounted for by Session. The reduction in the deviance statistic for between Model 1 and Model 2 for number of activities participated in was 965.65 and a reduction in BIC of 959.04. Reductions between Model 1 and Model 2 for all three measures were significant at the $p < .01$ level as judged by the critical chi-square statistic thus indicating that accounting for Session improves the model's fit to the data.

Model 3 introduced a random linear effect of Session and allowed the model to account for the data's covariance structure. Again, the AR(1) covariance structure was applied and resulted in a non-zero change in goodness-of-fit statistics. As such, Heck and colleagues' (2010) suggestions were followed and the unstructured covariance structure used for the proceeding growth models. Due to Model 3 including the same fixed effects as Model 2, the fixed estimates remain the same across both models. The variance components

identified in Model 2 developed for behavioural predictions (σ_{ε}^2 , σ_0^2 , σ_1^2 in Table 6.8) were significant ($p < .001$), meaning that clinicians are likely to have different intercepts and slopes in terms of their predictions of percentage of activities completed across 12 sessions of CBT. As such, future models should account for this by including random effects. The covariance between the true intercept and true rate of change (σ_{01}^2) yielded a significant positive outcome indicating that clinicians who predict a lower percentage of activities participated in at sessions 12, also predict a slower rate of progress over the course of therapy. The inclusion of random effects demonstrate an improved model fit based on statistically significant change in deviance and BIC values indicating a substantial change (difference > 10), further supporting the notion that accounting for differences between clinicians in predictions of final score and rate of change provides a better fit to the data.

The quadratic effect of time was modelled in Model 4, with the γ_{20} value indicating that the quadratic effect was significant for behavioural change with an overall true rate of -0.24. Behavioural improvement (i.e. percentage of activities completed) was predicted to decrease in magnitude from 8.39 to 3.02. Based on the γ_{00} produced in Table 6.8, as well as the value provided in the reverse Session analysis, behavioural change was expected to improve from 2.81 to 65.53. The drop in values between Model 3 and Model 4's σ_{ε}^2 value, along with the (R^2_{ε}) value indicate that taking into account the quadratic effect of time explains 36% of the variance of behavioural change predictions. In Table 6.8, the significant σ_0^2 , σ_1^2 , σ_2^2 , values indicate significant variability between clinicians in predictions of percentage of activities participated in at session 12 and the rate of change. The significant covariance value (σ_{01}^2) shows that clinicians who predict a lower percentage of activities

completed at session 12 typically predict a slower rate of change overall. Model 4 was deemed to be a better fit than the random effects linear growth model as judged by significant changes ($p < .05$) between the deviance statistics and a substantial change (i.e. > 10) in BIC statistics.

Incorporating the effect of Session³ as a fixed and random effect, Model 5 tested whether a cubic polynomial significantly improved the fit of the regression model for the percentage of activities Ms S completed. As can be seen in Table 6.8, γ_{30} demonstrates a non-significant effect ($p < .05$). This indicates that clinician's predictions of behavioural change is not best represented using a cubic trend and, as such, reverting to Model 4 is indicated as the best growth model to proceed in future analyses.

In the values provided for Model 6 in Table 6.8, the non-significance of the values γ_{01} (8.94, *ns*) and γ_{11} (0.82, *ns*) indicate experience level has little to no effect on the variability of predictions of behavioural change modelled on linear growth. The deviance statistic and BIC also demonstrated non-significant and non-substantial changes for percent of behaviours participated in, showing that modelling these effects did not provide a better fit to the data.

Applying the effect of experience to both linear and quadratic predictions again yielded non-significant values for γ_{01} (7.64, *ns*) and γ_{11} (1.71, *ns*), and the additional γ_{12} (0.09, *ns*) indicating a lack of effect for experience on behavioural change predictions.

As a supplement to the remaining unexplained variance, simple linear regression was used to identify the frequency of linear, quadratic, and cubic predictions of Ms S's completion of feared/avoided situations for each participant. Behavioural change predictions

across therapy demonstrated a linear progression by 30 (44%), a quadratic trend by 23 (34%), and cubic by 15 (7%).

Analysis 3: Presence of Discontinuous Patterns of Change

As with Mr T, an early rapid response was determined using the criteria provided by Ilardi and Craighead (1994). Overall 25 (37%) clinicians predicted an early rapid response when treating Ms S. Based on the Tang and DeRubeis' (1999b) modified criteria, no clinicians predicted that Ms S would experience a sudden gain. While participants predicted score changes between two sessions on the BAI that met the criteria for a 7 point drop between sessions that was 25% of the client's overall change, the third criterion was not met by any of the predictions. Overall 6 (9%) clinicians predicted an anxiety spike when treating Ms S.

Analysis 4: Gradual, Steady Change versus Variable Change

Graphs of the predicted scores were visually inspected to identify whether change was expected to be gradual and smooth or variable. Gradual and smooth change was identified as change that did not demonstrate any noticeable deviations from a linear or quadratic trend. Variable change was identified as any change that showed at least one noticeable deviation from a smooth progression. As such, cubic trends were automatically identified as a variable based on trajectories where change goes in two directions. Overall symptom change, as predicted by BAI scores were identified as smooth and gradual for 47 (69%), with 21 (31%) clinicians predicting variable change. This difference was identified as significant by a chi-square goodness of fit test, $\chi^2 (1) = 9.94$, $p = < 0.05$. Ms S's level of

anxiety, as predicted by PSWQ scores, was predicted to progress in a smooth and gradual manner by 36 (53%) clinicians and predicted to demonstrate variable change for 32 (47%) clinicians, with this difference deemed not significant ($\chi^2 (1) = 0.24$ $p = 0.628$). Behavioural change, as predicted by the percent of activities completed between sessions, was expected to be smooth and gradual by 46 (68%) clinicians, with 22 (32%) predicting variable change. Chi square goodness of fit analysis identified this difference as significant, $\chi^2 (1) = 8.47$, $p = < 0.05$

Analysis 5: The Relationship between Overall Symptom, Mood, and Behavioural Change

The number of clinicians who predicted each type of pattern can be seen in Table 6.9, on page 173.

Following on from this analysis, the type of change predicted to occur first for an anxious client was identified. Initial overall symptom occurring first was predicted by 3 (4%) clinicians and mood was expected to improve first by 10 (15%) Behaviour was expected to change first by more than half of the participants (36; 53%). Overall symptoms and mood were expected to simultaneously improve first by 3 (4%) clinicians, overall symptoms and behaviour were predicted to simultaneously improve first by 1 (2%) clinicians, and mood and behaviour were expected to simultaneously improve first by 7 clinicians.

Table 6.9.
Frequency of Each Sequence of Measured Dimension Expected for a Client with Anxiety

Order of change	Frequency
Simultaneous (SBM)	8
Overall symptom improvement then behavioural improvement then mood improvement (S-B-M)	2
Mood improvement then overall symptom improvement followed by behavioural improvement (M-S-B)	4
Mood improvement then behavioural improvement then overall symptom improvement (M-B-S)	3
Behavioural improvement then overall symptom improvement then mood improvement (B-S-M)	6
Behavioural improvement then mood improvement then overall symptom improvement (B-M-S)	8
Overall symptom improvement then simultaneous mood and behavioural improvement (S-MB)	1
Mood improvement then simultaneous overall symptom and behavioural improvement (M-SB)	3
Behavioural improvement then simultaneous overall symptom and mood improvement (B-SM)	22
Simultaneous overall symptom and mood improvement then behavioural improvement (SM-B)	3
Simultaneous overall symptom and behavioural improvement then mood improvement (SB-M)	1
Simultaneous mood and behavioural improvement then symptom improvement (MB-S)	7

Analysis 6: Clinically significant symptom reduction by the final session of therapy

When understanding the level of improvement of overall symptoms at session 12, the average BAI score as predicted by all clinicians was 21.65 (SD = 7.75). Use of Jacobson & Truax's clinical significance classification method indicated that 6 (9%) clinicians overall predicted scores at session 12 to reflect Improved change, with 62 (91%) predicted scores reflective of the Recovered range. The average score across clinicians for level of anxiety at

session 12, as predicted using the PSWQ, was 36.19 (SD = 11.82). Analysis of clinically significant change using the PSWQ indicated that 64 (94%) of all clinicians predicted scores reflecting the Improved range, and 4 (6%) predicting scores reflective of the Recovered range. None of the predictions made on either of these two measures indicated that Ms S would experience no change in terms of clinical significance classification.

Analysis of clinician predictions of the percentage of activities Ms S participated in the week prior to the final session indicated an average prediction of 66% (SD = 22.25). Of the predictions made, 18 (27%) clinicians predicted that Ms S would complete 50% or less of the activities she aimed to complete, and 50 (74%) predicted that she would complete over 51%. Furthermore, 50% of clinicians predicted Ms S to be completing 70% or more of her anxiety provoking situations at session 12.

Analysis 7: Maintenance of therapeutic change at 3- and 6-month follow-up

Table 6.10.

Significance of Mauchly's Test of Sphericity for Session 12, 3 Month Follow-up and 6 Month Follow-up for Overall Symptom Change (BAI), Level of Worry (PSWQ) and Percentage of Activities Participated In

Measure	Mauchly's W	Degrees of freedom	Significance
PSWQ	0.665	2	.000
BAI	0.528	2	.000
Percent of Activities	0.572	2	.000

Using a repeated-measures ANOVA design, significant differences in change between session 12, 3 month follow-up and 6 month follow-up were tested. For predictions of overall anxiety symptoms (BAI), level of worry (PSWQ), and percentage of activities participated in, Mauchly's test indicated that the assumption of sphericity was violated ($p <$

.05). As can be seen in Table 6.10, above, the Mauchly's test statistic for all three measures was under 0.75, the degrees of freedom were corrected using Greenhouse-Geisser estimates of sphericity.

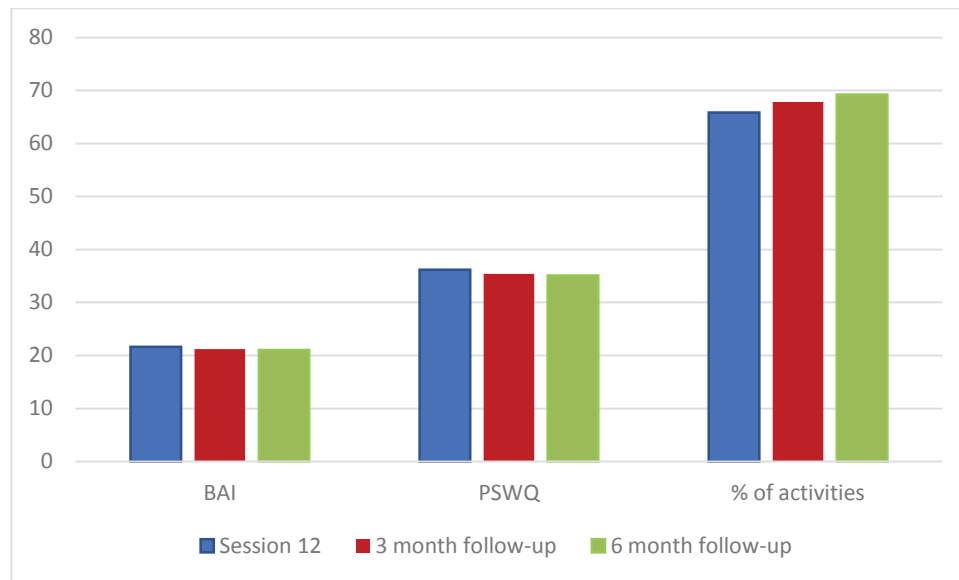


Figure 6.4. Average BAI, PSWQ and percentage of activities score for Session 12, 3 month follow-up and 6 month follow-up.

The average of predictions made on the BAI, PSWQ, and percentage of activities for Session 12, 3 month follow-up and 6 month follow-up are summarised in Figure 6-4, above. Results of the repeated measures ANOVA indicated that predictions in overall symptom change at session 12, 3 month follow-up and 6 month follow-up did not demonstrate significant differences, $F(1.36, 951.10) = 0.96$, $p = 0.36$, with predictions in worry level also demonstrating no significance in the amount of change between follow-up sessions, $F(1.50, 98.90) = 1.54$, $p = 0.22$. The percentage of activities Ms S was able to participate in demonstrated significant change across follow-up sessions, $F(1.40, 92.41) = 5.80$, $p = .01$. As the assumption of sphericity was violated for the percentage of activities, Bonferroni's

adjustment was applied when making pairwise comparisons (Field, 2009). This post-hoc testing indicated that there was a significant difference between session 12 and 6 month follow-up.

As with Mr T, the frequency of deterioration and improvement, alongside the magnitude of such change was addressed. As experience was determined to have no impact on predictions, this was not applied as a within subjects factor. Beginning with change between session 12 and 3 month follow-up, Ms S's overall symptom change was expected to demonstrate no change by 8 (12%) clinicians, improve on average 3.09 points, with a maximum improvement of 8 points by 24 (50%) clinicians, and deteriorate on average by 2.88 points, with a maximum prediction of 17, by 26 (39%) clinicians. Level of worry, determined by the PSWQ, was predicted to improve by 34 (50%) clinicians, at an average rate of 4.97 point with a maximum of 16, with 27 (40%) clinicians predicting deterioration at an average of 4.26 points with a maximum of 13. 7 (10%) clinicians predicting no change to level of worry. Ms S's ability to participate in activities was expected to improve by 36 (53%) clinicians, deteriorate by 21 (31%) clinicians, and stay constant by 11 (16%) clinicians. The average amount of improvement was determined to be 8.11% with a maximum of 22, and average amount of deterioration being a reduction of 7% with a maximum of 21.

Between 3 and 6 month follow-up, 20 (29%) clinicians predicted that Ms S would demonstrate no change in overall symptoms. On average, those (27; 43%) who identified improved scores for Mr S over this time period indicated an improvement of 2 in BAI score, with a maximum of 12 points. Those who predicted deterioration (17; 25%) expected an

average of 2.94 points reduction with the maximum amount of deterioration being 8. Ms S's level of worry was expected to improve by 33 (49%) clinicians with an average improvement of 2.67 point on the PSWQ, and a maximum of 13. 18 (27%) clinicians predicted Ms S would demonstrate an on average 3.78 point deterioration during this time period (maximum deterioration: -10 points), with 17 (25%) clinicians expecting no change. The percentage of activities that Ms S was effectively participating in between 3 and 6 month follow-up was predicted to remain constant by 19 (28%) clinicians, improve with an average of 5% by 29 (43%) clinicians, and deteriorate with an average of 3% by 20 (29%) clinicians. The maximum amount of improvement predicted was 19, and the most deterioration predicted was a 12% drop in activity completion.

For predictions in regards to overall symptoms, BAI scores were expected to show no change overall between session 12 and the 6 month follow-up by 5 (7%) clinicians. Improvement, at an average level of improvement of 3.58 points, was predicted by 38 (56%) clinicians. Deterioration, at an average level or improvement of 3.76 points, was predicted by 25 (38%) clinicians. Level of anxiety, measured by the PSWQ, was predicted to stay consistent over this six month time period by 7 (10%) clinicians. Improvement was expected by 38 (56%) clinicians, at an average score of 5.16 points, and deterioration expected by 23 (34%) clinicians, at an average score of 5.30. The percent of activities Ms S participated in between the final session and six month follow-up were expected to improve, on average by 8.90, by 41 (60%) clinicians, and expected to deteriorate by 18 (27%) clinicians, by an average of 7.61. No change was expected on the percent of activities Ms S participated in between this time period by 9 (13%) clinicians.

Clinically significant changes were assessed over both the Session 12 to 3 month period, the 3 month to six month period and overall between session 12 and 6 month follow-up for BAI and PSWQ predictions. Results based on changes in overall symptoms (i.e. BAI scores) showed that 62 (91%) clinicians predicted that Ms S would show no change in her clinically significant classification predictions. Three clinicians predicted improvement from Improved to Recovered, with another 3 prediction a shift from Recovered to Improved. Completing the same form of analysis for PSWQ scores, which addressed level of anxiety, showed that 63 (93%) clinicians predicted no change in clinical significance classification. One clinician predicted a shift between session 12 and 3 month follow-up from Improved to Unchanged, 2 predicted a shift from Improved to Recovered, and 2 clinicians predicted a shift from Recovered to Improved.

When identifying shifts in clinically significant classification between the period of 3 month and 6 month follow-up for BAI scores, Ms S was predicted to demonstrate no change by 64 (94%) clinicians, with 1 clinician predicting a shift from Improved to Unchanged, 2 clinicians predicting a shift from Improved to Recovered and 1 clinician predicting a shift from Recovered to Improved. PSWQ scores were expected to remain Unchanged over this time period by 65 (96%) clinicians, with 3 clinicians predicting a shift from Improved to Recovered.

Session 12 to 6 month follow-up analysis revealed that 61 (90%) clinicians predicted no change in clinically significant classification when considering overall symptoms, predicted using BAI scores. Based on the scores at session 12 and 6 month follow-up, 4 clinicians expected Ms S's classification to change from Improved to Recovered, 1 clinician

expected a change from Recovered to Unchanged and 2 clinicians expected a change from Recovered to Improved. Level of anxiety, as measured by PSWQ scores, were expected to demonstrate no change in clinical significance classification by 62 (91%) clinicians, with 1 clinician expecting a shift from Improved to Unchanged, 4 clinicians expecting a shift from Recovered to Unchanged, and 2 clinicians predicting a shift from Recovered to Improved.

CHAPTER SEVEN: DISCUSSION

This chapter presents the main findings of the study in regards to the research questions posed in Chapter One. Following this discussion, the contributions this study makes to the research literature will be presented. This will then be followed by limitations, implications, and future directions which are apparent as a result of this study.

Summary of Key Findings

Clinician awareness of theories of change

The purpose of this question was to act as a primer for understanding what influences may have impacted on clinicians' predictions of change across the course of therapy. Based on the results of the current study, it appears that few clinicians (15; 22%) demonstrated an awareness of at least one change model. In terms of the change models highlighted in the literature, the stages of change model, the three phase model, and the theory of sudden gains were mentioned. This is likely to be reflective of both the emphasis of therapy outcome at the end of therapy, rather than the process, as well as a lack of definition.

Pattern of change across therapy between session 2 and 12

While the majority of prior studies identified change progression in terms of a single measure, with those measures typically being overall symptoms such as the BDI-II and BAI, the present study identified the individual expected trends of change for three measures concerning symptoms, mood, and behaviour separately. Clinician predictions indicated that, averaged across all clinicians, overall symptoms, mood, and behaviour followed a quadratic trend for both a depressed and anxious client with the magnitude of change predicted during

the initial sessions more rapid than the magnitude of change as the end of therapy nears. A quadratic trend is one that displays concavity, a single bend either upward or downward. In the context of this study, such a pattern reflects a decelerating curvilinear trend reflecting Howard and colleagues findings (Howard et al., 1986) that indicates that clients with depression or anxiety are likely to experience considerable change early in therapy with progress levelling out as therapy continues. In addition to MLM indicating that the quadratic trend best encapsulates how change progresses across therapy, it further indicates that there was significant variability between clinicians in terms of their predicted trends. This thus illustrates that there is no overarching cohesive expectation of the slope or rate of change for client change that encompasses what clinicians expect. As such, further regression analyses were used to explore the number of clinicians who predicted a linear, quadratic, or cubic trend.

For the client with depressive symptoms, analysis of overall symptom change (BDI-II scores) indicated 50% of clinicians predicted a linear trend, 41% predicted a quadratic trend and 9% predicted a cubic trend, where plotted progress changes in direction (e.g. from improving to deteriorating) on two occasions. A similar finding is found for behavioural change with a higher number of clinicians ($N = 38$; 56%) predicting a linear trend, with 21 (31%) predicting a quadratic trend and 7 (10%) predicting a cubic trend. Predictions of negative affect (PANAS-NA scores) demonstrated a higher rate of quadratic trends ($N = 27$; 40%) when compared with linear predictions ($N = 23$; 34%) and cubic trends ($N = 17$; 25%). While the number of clinicians who predict a quadratic trend is not significantly smaller than those who predict a linear trend for both overall symptom and behavioural change, there is a large number of clinicians, at times larger than the number of clinicians who predicted a

quadratic curve, who predict a linear trend. This shows that a number of clinicians predict in line with the literature's established decelerating curvilinear trend of change. Despite more clinicians predicting a quadratic trend as opposed to a linear trend for changes in negative affect, only a small difference is apparent, indicating that a substantial number of clinicians who were participants in this study predicted that client change would be linear (i.e. a steady, straight line progression from the start of therapy to its end) despite the literature's expectation that the majority of clinicians would predict a quadratic trend.

When considering a client receiving CBT for the treatment of anxiety, again, overall symptom change (BAI), level of anxiety (PSWQ), and behavioural change (percentage of activities completed) analyses indicated that across the three measures, a higher number of clinicians predicted a linear trend (50%, 47%, and 44% respectively). A quadratic trend was predicted by approximately a third of clinicians across all three measures, with such a trend predicted by 32% of clinicians for overall symptoms, 29% for level of anxiety, and 34% for percent of activities. The cubic trend was predicted by 16% when considering overall symptoms, 24% for level of anxiety and 22% for percentage of activities participated in. As with overall symptom and behavioural change for a client exhibiting depressive symptoms, these results show that more clinicians predict a linear trend, which is in contrast to the decelerating curvilinear trend demonstrated in the session-by-session research literature.

When the data were analysed using MLM statistics, they showed that the pattern of change predicted by the participants were commensurate with findings from the currently available literature. However, further exploration revealed that more clinicians appear to predict linear change, reflective of a trend with little empirical support. As such, this finding

is indicative of the frequently claimed gap between scientific knowledge and clinical assumptions. Of course it remains uncertain as to whether the clinicians are unaware of the current literature or simply have different experiences within their own practice.

Differences between experienced and trainee clinicians

The effect of experience was explored in the current study to identify whether or not there were differences in how change was understood. No significant effect was identified between experienced and trainee clinician predictions for both the client exhibiting Depression and the client exhibiting Generalised Anxiety Disorder. This shows that experience level had no impact on the type of trend a clinician predicted. Furthermore, awareness of relevant change theories was shown to be not dependant on clinician experience, however, experienced clinicians were able to identify other theories that they attributed to their understanding of change.

Prediction of discontinuous patterns

When I defined early rapid response using Ilardi and Craighead's criterion of early rapid response, I found that 13% of experienced and training clinicians predicted such a discontinuous pattern in the treatment of Mr T. Despite prior research indicating that between 17 to 39% of clients experience a sudden gain as defined by the criteria set by Tang and Derubeis (1999; Stiles et al., 2003), no clinicians in the current study predicted a sudden gain occurring at any time across the course of therapy for Mr T. Only 3 (4%) clinicians predicted a depression spike.

Approximately a third (37%) of clinicians predicted an early rapid response for the anxious client. As in the case of their expectations for Mr T, no clinicians predicted a sudden gain for Ms S. An anxiety spike was expected by 6 (9%) of the clinicians.

Taken alongside clinician awareness of theories of change, and the high number of clinicians who predicted a linear trend, I think that the reason for this finding is that the participants were not aware of the literature on discontinuous patterns, consequently they do not expect either depressed or anxious clients to demonstrate substantive changes in rate of change. Again, this finding is indicative of the frequently claimed gap between science and practice. Furthermore, while the number of clinicians who identified an early rapid response for Ms S may indicate that some clinicians are aware of a rapid early response in the treatment of clients with anxiety, I also feel that it shows that clinicians expect change to be different between depressed and anxious clients in terms of overall symptom change, with anxious clients demonstrating a quicker rate of improvement.

Smooth, gradual change versus variable change

The literature on dynamical systems theory (Hayes et al., 2007b) proposes that change can be conceptualised as smooth and continuous or variable and discontinuous, with variable change reflecting the pattern that is most typically seen in clinical practice. In terms of the research on general variability across the course of therapy, however, no formal understanding of this theory has been explored aside from discontinuous patterns. This is likely due to exploration of individual patterns being difficult statistically. As such, the current study took the opportunity to identify how change is conceptualised by the experienced and trainee clinicians through the visual inspection of graphs. In terms of a client

with depression, overall symptom change was expected to be predominantly smooth and continuous with 65% of participants predicting such a pattern. In contrast, negative affect and behavioural change were predicted to be relatively evenly split between clinicians, with no predominantly smooth and continuous or variable and discontinuous pattern of change identified. For overall symptom and behavioural change in an anxious client, clinicians did not predict a single predominant pattern of change. Level of anxiety was expected to demonstrate a predominantly smooth and continuous change pattern across the course of therapy. Again, this tendency to report smooth and gradual change indicates a possible lack of continuity between science and practice.

It is important to acknowledge that the current study's definition of smooth, continuous change and variable, discontinuous change is subject to researcher interpretation and, as such, further exploration of these two types of patterns is likely to be improved by comprehensive definitions of both smooth, gradual change and variable, discontinuous change. The definitions used in my study could be considered arbitrary as they could be substituted with other, more comprehensive and easily utilised definitions which would possibly provide different results. This could then allow evaluation of the impact of either smooth or variable change on a client's overall outcome or rate of change. The overall outcome of this exploration is, however, that clinicians do predict both smooth and variable trends of overall symptom, mood, and behavioural change, with smooth overall being predominant. This again may be interpreted as being in contrast to the research literature on discontinuous patterns, where variability has been identified.

Relationship between symptomatic, mood, and behaviour change

As emphasised throughout this research report, an understanding of what forms of change occur prior to other forms of change, in this case overall symptoms, mood, and behaviour, has not been established. There have been hints at potential patterns, such as within CBT circles where it has been suggested that a client with depression is likely to demonstrate changes in their behaviour prior to improvement in mood and overall symptoms (Beck, 2011; Cuijpers et al., 2007; Hollon & Beck, 2013; Lejuez et al., 2001; Samoilov & Goldfried, 2000). Clients with an anxiety diagnosis undertaking CBT have been previously hypothesised to show improvement initially in their mood, before engaging in behaviours and demonstrating overall symptom improvement (Evans, 2013; Wells, 1997). In terms of patterns of change based on predictions made by experienced and trainee clinicians, a prior study completed by Fletcher (Fletcher, 2011) demonstrated that for both a depressed and an anxious client, participants expected overall symptom and mood improvement to occur simultaneously, with behavioural improvement occurring initially. For the most part, these potential patterns have been derived from discussions in the research literature, without any direct research being conducted into understanding the exact nature of the patterns in which individuals demonstrate change throughout therapy. As such, my study provides an insight into what is expected, which may offer an understanding of what may be occurring in therapy.

When considering the pattern of change for overall symptoms, mood, and behaviour for a client undertaking CBT for the treatment of depression, 19 clinicians predicted that improvement in behaviour would occur first, followed by improvement in mood, and overall symptoms expected to improve last. As such, clinicians expected that a client should begin

engaging in activities before expecting mood improvement, and expected improvement in behaviour and mood to occur before overall symptoms begin demonstrating a large amount of improvement. This appears to be in line with suggestions of patterns within the CBT literature. When considering singular change predictions and change where two types of change occurred simultaneously, 50 clinicians expected behaviour to be one of the first types of change to improve, further supporting the notion that behaviour should show improvement initially prior to a client demonstrating decreases in their presenting symptoms.

Another pattern expected by a number of clinicians (17; 25%) involved initial simultaneous improvement in mood and behaviour, followed by an improvement in overall symptoms. Simultaneous improvement, where no single type of change occurred before the others; was also a common prediction, with 13 clinicians expecting this. These patterns thus provide us with an alternative insight into the pattern of change that is expected to occur for a client with depression, and is likely to guide further explorations into the pathways of change for a client with depression.

When considering a client who is undertaking CBT for anxiety, the dominant pattern predicted was improvement in the client's participation in activities previously avoided, followed by simultaneous overall symptom and mood improvement. When looking at the type of change expected to manifest first, behaviour made up 44 of the singular and simultaneous predictions, again suggesting that resisting avoidance was expected to improve prior to mood and overall symptom improvement. This differs from the suggestion derived from the CBT literature that clients are likely to show improvement in their mood first,

before demonstrating improvements in their behaviours and demonstrating overall symptom improvement.

Level of change by session 12

Using the frequently utilised Jacobson and Truax method of calculating clinical significance (Jacobson & Truax, 1991), it was found in my study that clinicians predict both a depressed client and an anxious client to demonstrate symptom recovery by the end of a 12 session course of CBT. Mood change, depicted by negative affect for the depressed client and level of anxiety for the anxious client, was predicted to demonstrate Improvement by the majority of clinicians.

What defines clinically significant behaviour change is yet to be established (although interest in developing norms and evaluating reliability and validity, see Craske & Tsao (1999)), and the success of number of activities participated in, or percentage of anxiety provoking situations completed, is dependent on the client's goals and requirements of functioning. A limitation of the current study is that it did not identify a goal and, if it had, this could have been used to identify how successful a depressed or anxious a client is expected to be in terms of his or her behavioural change. Despite this, the majority of clinicians expected a depressed client to participate in six or more enjoyable activities indicating a 60% improvement from the zero level specified in the pre-treatment baseline session. For the anxious client, another limitation of this study was that the graph plotting programme was not divided into 11 units (as the electronic task indicated 11 feared situations). This meant that clinicians predicted variable completions of task, which may in actual fact be reflective of anxiety work, where attempts may be counted. However, there is

no way to establish this in the current study. Despite this limitation the majority of clinicians predicted over 50% of anxiety provoking situations completed, with at least half of clinicians predicting Ms S to expose herself to 70% or more feared situations. As such, definite change is expected for both the depressed and anxious client, although not to the extent of a 100% success rate.

In terms of clinician confidence, New Zealand clinicians were shown to demonstrate confidence in their expectations of change at Session 12 (i.e. the final session of therapy). This can be interpreted in a number of ways, with clinicians either having confidence in their own abilities, confidence in their use of CBT, and/or confidence in the treatment of depression or anxiety with CBT.

With the CBT literature offering brief methods of therapy (Beck, 2011), along with symptom measure having a higher likelihood of being familiar to clinicians (Froyd, 1996; Mellor-Clark et al., 1997; Society for Psychotherapy Research, 1993), it is no surprise that clinicians expect depressed and anxious clients to recover in the 12 sessions, despite Hansen and Lambert (Hansen & Lambert, 2003) concluding that positive therapeutic change is typically only achieved after 14 sessions. The expectation that mood/anxiety and behavioural change is not completely resolved by the final session (i.e. majority of clinicians predicted a clinically significant classification of “improved”) can be interpreted as meaning that clinicians do not predict that client problems are completely resolved after 12 sessions of CBT, consistent with the literature (Hansen & Lambert, 2003). Although no clinicians predicted that either client in the present study would deteriorate, further research is required to establish whether this is due to the presentation of only two clients with different mental

health issues, or clinicians' excessive confidence in themselves. Future research in the area of clinician optimism or awareness of deterioration may benefit from the inclusion of a number of client cases as well as the inclusion of predictions that consider both improvement and deterioration. Taken alongside the predictions of statistically meaningful change, the current study implies that clinicians appear confident, albeit not to the extent of excess, in the outcome of their clinical therapy. This appears to be further supplemented by clinicians predicting that client outcome is maintained, as opposed to continuing to improve significantly following therapy.

Maintenance of therapeutic change at 3 and 6 month follow-up

Findings of the present study indicated that the majority of clinicians expected both a depressed client and an anxious client to demonstrate little change in scores and expected participation in activities over three and six month follow-up. This indicates that clinicians expect clinical change to be maintained following the cessation of therapy sessions, as would be expected based on the literature. This alludes to confidence in clinician clinical skills and perhaps in CBT as a therapeutic modality.

Clinical and Practical Implications

This study has focused on the patterns of change predicted by clinicians and how these predictions compare with the current change research literature. The main finding in regards to this is that clinicians' expectations do not appear to follow what is identified in the research literature thus demonstrating a possible presence of the scientist-practitioner gap in New Zealand. Whether or not this can be attributed to lack of awareness of the research literature, or is reflective of the true nature of clinical practice, still requires further

exploration. When considering whether or not level of experience impacts on predictions of change in therapy, the present study identified similarities in predictions between experienced clinicians and trainees, suggesting that this current understanding of change is not solely based on experience within the clinical setting, but also on what trainees are being exposed to in their clinical training.

In more recent years psychotherapy research has progressed into understanding how clients change, to enhance the application of empirically backed modalities and techniques by understanding at what stage of therapy aspects of change are best targeted. In addition, understanding how clients change, further prepares clinicians to identify those clients who are not progressing in therapy as well as they should. This is important as being aware of what should be expected of client change can assist in identifying when a client requires further intervention in a timely manner so as to avoid attrition or a negative outcome. Understanding change also allows a better understanding of prognosis based on the client's performance in the earlier therapeutic sessions. As such, my findings in the current investigation lead me to consider that there may be an opportunity for both trainee and experienced clinicians to develop their understanding of what is expected of change across the course of therapy, as a way to improve the application of an empirically based treatment as well as enabling the identification of clients who may require additional intervention. The method used in this study also holds potential as a training method to develop such an understanding, with the possibility of training clinicians to develop their understanding of what to expect of change by comparing their predictions to hypothetical outcomes for clients that reflect more precisely what has been found in the treatment literature. This may be enhanced by having real data for numerous clients to demonstrate possible trajectories and

the likely outcomes that result based on initial change or the presence of discontinuous patterns. This will thus prime future clinicians for what to expect well before they experience their first therapeutic relationship.

The present study also offers further understanding of how routine outcome measurement can be implemented as a way to monitor client progress and address any issues that may arise. Clinicians in this study demonstrated having an understanding of what they expect to happen in therapy, therefore, deviations from these predictions are likely to alert clinicians to factors that may suggest a client is not progressing as well as they should, or may require further intervention. As such, a case can be advocated for the inclusion of at least one form of regular measurement of outcome, alongside a clinician's own judgement, to identify whether clients are improving in the way that they expect, or whether alternative intervention plans are required.

This study continues on from patient-focused research, demonstrating strengths associated with using this methodology. While this study showed what type of changes can occur throughout therapy, it also allows for an understanding of the expected effectiveness of CBT and clinician confidence in their ability to utilise CBT in clinical practice. Additionally, this study provides an opportunity for a more in depth analysis of the type of change that takes place at what point in therapy and why those who deviate from the expected path, or drop out, do so. Applying the method used in understanding change expectations in this study, it is also possible to explore further in terms of the individual characteristics of a client. Potential future avenues of this research include assessing client factors such as age, gender, ethnicity, type of illness, and symptom severity. Further expansion would include comparisons between psychologists in other countries as well as modes of therapy. Furthermore, a study design such as this can expand the

knowledge surrounding such theoretical concepts as mechanisms of change and common factors with clinician expectations again being explored through the use of case studies.

Limitations

Following the completion of this study, a number of limitations were identified. If it could be done again, a larger sample would be preferred for this particular methodology. With a longer allowance of time for participants to complete tasks and wider distribution, in particular to experienced clinicians, it would ensure that an adequate representation of expectations and views are obtained.

The use of hypothetical clients may act as a limitation to what can be inferred from this study's results, particularly as progress in therapy for clients fulfilling the descriptions provided is currently unanswered. While the purpose of the study was to look at what clinicians expect as opposed to how a client will change, it would be beneficial to have some frame of reference which can be compared to predictions. Without the ability to compare the expectations with real life data, it limits the comparison between research and experience due to the literature reflecting one perspective (i.e. the research perspective) and experienced clinicians reflecting a different perspective (i.e. the clinical practice perspective). As such, if further studies were to be conducted using this methodology, I would recommend that a number of real cases be observed to provide a statistically founded case composite that provides a basis to compare with both the research literature and clinician predictions. Clarification around whether these case composites represent an average, improving, or deteriorating client in future studies would determine more comprehensive understanding of what clinicians are basing their predictions on.

The case vignettes could have been improved to emphasise the principles of CBT by incorporating a case conceptualisation, more of an emphasis on cognitive change and clarification around the therapeutic alliance. The incorporation of a case conceptualisation would firmly ground the presented cases in cognitive behavioural theory and provide features similar to what a typical CBT clinician would expect when planning treatment. A case conceptualisation would further allude to cognitions that may provide further ideas on the predictions that may be expected for particular clients. Future explorations in this area should also include a cognitive change measure to comprehensively understand the sequence of change in CBT. While the therapeutic alliance was alluded to, it has been acknowledged as a primary principle in CBT and, as such, this therapeutic factor could have been highlighted more clearly.

This study's results could have been improved with the use of a behaviour psychometric, to allow for the identification of clinically significant change for all three measures, or specification of the client's end of treatment goal, for a point of comparison. This would have also allowed for a better point of comparison in terms of the follow-up sessions as opposed to solely using improvement and deterioration without specific parameters. While the average predicted score on the behavioural measure was provided, it does not indicate how meaningful this change is statistically, or for the client themselves.

Although clinicians were asked if they were aware of none, some, or all measures, a better understanding could have been acquired from asking specifically which measures clinicians were aware of. This may have identified predictions made with stronger confidence by clinicians, and thus improving the validity of predictions, especially in terms of trainee clinicians who indicated less awareness of all measures than experienced clinicians. Being familiar with the measures used would also supplement the understanding

of indications of meaningful change. A clinician who is more familiar with a particular measure may also have an increased awareness of what score on that measure reflects a clinically significant outcome. As such, interpretation of the overall predictions of change must be done with caution as inexperience with a measure may confound some of the predictions.

Asking clinicians about how regularly they measure change across the course of therapy, and with what measure, would have broadened the understanding of clinician predictions. Clinicians who regularly implement routine outcome measurement may have shown differences in their trends based on their increased experience in observing how clients actually change across therapy. As such, future research in this area is likely to benefit from including this approach.

Conclusion

The primary aims of this study were to explore what patterns of change for both a depressed and anxious client undertaking 12 sessions of CBT would be expected by both experienced and novice clinical psychologists familiar with CBT. The research question was further divided into (a) the overall pattern of change expected, (b) the presence of discontinuous patterns, (c) whether change is expected to be smooth and gradual overall or variable, (d) how overall symptoms, mood, and behavioural change interrelate, (e) the level of outcome expected at the end of therapy, and (f) whether or not this outcome is maintained.

Sixty eight clinicians, with exactly half being experienced clinicians and the other half being trainee clinicians, were invited to complete the electronic task used in the current

research. The study showed that clinician expectations of change do not appear to adhere to the research literature and are not dependent on level of experience. This may indicate that there is a mismatch between research and clinical practice that continues to require resolution. It further indicates possible training opportunities, especially in terms of preparing future clinicians by increasing insight into the types of change that can occur across the course of therapy to promote understanding of what is likely to lead to a positive outcome, as well as how to identify when a client is not progressing as they should. The present study further shows that clinicians appear confident in their ability to achieve positive outcomes using CBT and believe CBT is effective in the treatment of depression and anxiety in the provided time frame. This suggests that not only does CBT have the support of empirical research; it is also viewed as a helpful modality by clinicians. As such, it appears to continue to be a good base modality to train clinical psychologists in New Zealand. In addition, New Zealand clinicians do not appear excessively confident in their ability to achieve positive outcomes, suggesting that they may be in a better position to be aware of clients who are deteriorating as opposed to clinicians who have been identified as overly optimistic. As such, the present study provides a useful starting point to build on in terms of understanding client change when considering the New Zealand context, and indeed, generalizing the approach, at a future date, into a multi-site, multi-national study to extend our understandings as change agents in the field of psychotherapy.

REFERENCES

- Aas, R. W., & Alexanderson, K. (2012). Challenging evidence-based decision-making: A hypothetical case study about return to work. *Occupational Therapy International*, *19*(1), 28-44.
- American Psychiatric Association (Ed.). (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). Arlington, VA: American Psychiatric Association.
- American Psychological Association. (1987). *Diagnostic and statistical manual of mental disorders*. Washington, DC: American Psychological Association.
- American Psychological Association. (2007). *APA dictionary of psychology*. Washington, DC: American Psychological Association
- APA Task Force on Evidence Based Practice. (2006). Evidence-based practice in psychology. *American Psychologist*, *61*(4), 271-285.
- Bandura, A. (1997). *Self-efficacy: The exercise of self-control*. New York, NY: Freeman.
- Barkham, M., Connell, J., Stiles, W. B., Miles, J. N., Margison, F., Evans, C., & Mellor-Clark, J. (2006). Dose-effect relations and responsive regulation of treatment duration: The good enough level. *Journal of Consulting and Clinical Psychology*, *74*(1), 160-167.

- Barkham, M., Evans, C., Margison, F., McGrath, G., Mellor-Clark, J., Milne, D., & Connell, J. (1998). The rationale for developing and implementing core outcome batteries for routine use in service. *Journal of Mental Health, 7*(1), 35-47.
- Barkham, M., Rees, A., Stiles, W. B., Shapiro, D. A., Hardy, G. E., & Reynolds, S. (1996a). Dose-effect relations in time-limited psychotherapy for depression. *Journal of Consulting and Clinical Psychology, 64*(5), 927-935.
- Barkham, M., Stiles, W. B., Hardy, G. E., & Field, S. (1996b). The assimilation model: Theory, research and practical guidelines. In W. Dyden (Ed.), *Research in counselling and psychotherapy: Practical applications* (pp. 1–24). London, UK: Sage.
- Barkham, M., Stiles, W. B., & Shapiro, D. A. (1993). The shape of change in psychotherapy: longitudinal assessment of personal problems. *Journal of Consulting and Clinical Psychology, 61*(4), 667-677.
- Barlow, D. H., Hayes, S. C., & Nelson-Gray, R. O. (1984). *The scientist practitioner: Research and accountability in clinical and educational settings*. New York, NY: Pergamon.
- Bauer, S., Lambert, M. J., & Nielsen, S. L. (2004). Clinical significance methods: A comparison of statistical techniques. *Journal of Personality Assessment, 82*(1), 60-70.
- Baumeister, R. F. (1991). *Meanings of life*. New York, NY: Guilford.

- Baxter, J., Kingi, T. K., Tapsell, R., Durie, M., & McGee, M. A. (2006). Prevalence of mental disorders among Māori in Te Rau Hinengaro: The New Zealand mental health survey. *Australian and New Zealand journal of Psychiatry, 40*(10), 914-923.
- Beck, A. T., Epstein, N., Brown, G., & Steer, R. A. (1988). An inventory for measuring clinical anxiety: psychometric properties. *Journal of Consulting and Clinical Psychology, 56*(6), 893-897.
- Beck, A. T., & Steer, R. A. (1990). *Manual for the Beck Anxiety Inventory*. San Antonio, TX: Psychological Corporation
- Beck, A. T., Steer, R. A., & Brown, G. K. (1996). *Manual for the Beck Depression Inventory-II*. San Antonio, TX: Psychological Corporation.
- Beck, A. T., Ward, C. H., Mendelson, M., Mock, J., & Erbaugh, J. (1961). An inventory for measuring depression. *Archives of General Psychiatry, 4*(6), 561-571.
- Beck, J. S. (2011). *Cognitive behavior therapy : Basics and beyond* (2nd ed.). New York: Guilford.
- Bickman, L., Rosof-Williams, J., Salzer, M. S., & Summerfelt, W. (2000). What information do clinicians value for monitoring adolescent client progress and outcomes? *Professional Psychology: Research and Practice, 31*(1), 70-74.
- Bishop, M. J., Bybee, T. S., Lambert, M. J., Burlingame, G. M., Wells, M. G., & Poppleton, L. E. (2005). Accuracy of a rationally derived method for identifying treatment

- failure in children and adolescents. *Journal of Child and Family Studies*, 14(2), 207-222.
- Bohart, A., & Tallman, K. (2010). Clients: The neglected common factor in psychotherapy. In B. L. Duncan, S. D. Miller, B. E. Wampold, & M. A. Hubble (Eds.), *The heart and soul of change: Delivering what works in therapy* (2nd ed., pp. 83-111). Washington, DC: American Psychological Association.
- Borkovec, T. D., Echemendia, R. J., Ragusea, S. A., & Ruiz, M. (2001). The Pennsylvania Practice Research Network and future possibilities for clinically meaningful and scientifically rigorous psychotherapy effectiveness research. *Clinical Psychology: Science and Practice*, 8(2), 155-167.
- Brinegar, M. G., Salvi, L. M., & Stiles, W. B. (2008). The case of Lisa and the assimilation model: The interrelatedness of problematic voices. *Psychotherapy Research*, 18(6), 657-666.
- Brown, J., Dreis, S., & Nace, D. K. (1999). What really makes a difference in psychotherapy outcome? Why does managed care want to know? In M. A. Hubble, B. L. Duncan, & S. D. Miller (Eds.), *The heart and soul of change: What works in therapy*. (pp. 386 - 406). Washington, DC: American Psychological Association.
- Burns, D. D., & Spangler, D. L. (2000). Does psychotherapy homework lead to improvements in depression in cognitive-behavioral therapy or does improvement lead to increased homework compliance? *Journal of Consulting and Clinical Psychology*, 68(1), 46-56.

- Bybee, T. S., Lambert, M. J., & Eggett, D. (2007). Curves of expected recovery and their predictive validity for identifying treatment failure. *Tijdschrift voor Psychotherapie*, 33(6), 272-281.
- Callahan, J. L., Swift, J. K., & Hynan, M. T. (2006). Test of the phase model of psychotherapy in a training clinic. *Psychological Services*, 3(2), 129-136.
- Campbell-Sills, L., & Brown, T. A. (2010). Generalized anxiety disorder. In M. A. Antony, & D. H. Barlow (Eds.), *Handbook of assessment and treatment planning* (2nd ed., pp. 224-266). New York, NY: Guilford.
- Carey, T. A., Carey, M., Mullan, R. J., Murray, L. K., & Spratt, M. B. (2006). Psychological change: What changes and how does it occur? A critical review. *Counselling Psychology Review*, 21(4), 28-38.
- Carey, T. A., Carey, M., Stalker, K., Mullan, R. J., Murray, L. K., & Spratt, M. B. (2007). Psychological change from the inside looking out: A qualitative investigation. *Counselling and Psychotherapy Research*, 7(3), 178-187.
- Caro Gabalda, I. (2006). The assimilation of problematic experiences in the context of a therapeutic failure. *Psychotherapy Research*, 16(4), 436-452.
- Cartwright, D. S. (1955). Success in psychotherapy as a function of certain actuarial variables. *Journal of Consulting Psychology*, 19(5), 357-363.
- Chambless, D. L., & Hollon, S. D. (1998). Defining empirically supported therapies. *Journal of Consulting and Clinical Psychology*, 66(1), 7 - 18.

- Christensen, L., & Mendoza, J. (1986). A method of assessing change in a single subject: An alteration of the RC index. *Behavior Therapy, 17*(3), 305-308.
- Clark, D. A., Hollifield, M., Leahy, R., & Beck, J. S. (2009). Theory of cognitive therapy. In G. O. Gabbard (Ed.), *Textbook of psychotherapeutic treatments* (pp. 165-200). Arlington, VA: American Psychiatric Publishing.
- Consumer Reports. (1995). *Mental health: does therapy help?*
- Craske, M. G., & Tsao, J. C. (1999). Self-monitoring with panic and anxiety disorders. *Psychological assessment, 11*(4), 466.
- Crawford, J. R., & Henry, J. D. (2004). The Positive and Negative Affect Schedule (PANAS): Construct validity, measurement properties and normative data in a large non-clinical sample. *British Journal of Clinical Psychology, 43*(3), 245-265.
- Crits-Christoph, P., Gibbons, M. B. C., & Mukherjee, D. (2013). Psychotherapy process-outcome research. In M. J. Lambert (Ed.), *Bergin and Garfield's handbook of psychotherapy and behavior change*. Hoboken, NJ: Wiley.
- Cuijpers, P., Berking, M., Andersson, G., Quigley, L., Kleiboer, A., & Dobson, K. S. (2013). A meta-analysis of cognitive-behavioural therapy for adult depression, alone and in comparison with other treatments. *Canadian Journal of Psychiatry, 58*(7), 376 - 385.
- Cuijpers, P., Van Straten, A., & Warmerdam, L. (2007). Behavioral activation treatments of depression: A meta-analysis. *Clinical Psychology Review, 27*(3), 318-326.

- Dawes, R. M. (1996). *House of cards: Psychology and psychotherapy built on a myth*. New York, NY: Free Press.
- Doss, B. D. (2004). Changing the way we study change in psychotherapy. *Clinical Psychology: Science and Practice*, *11*(4), 368-386.
- Dozois, D. J. A., Dobson, K. S., & Ahnberg, J. L. (1998). A psychometric evaluation of the Beck Depression Inventory–II. *Psychological Assessment*, *10*(2), 83-89.
- Duncan, B. L. (Ed.). (2010). *The heart & soul of change : delivering what works in therapy* (2nd ed.). Washington, DC: American Psychological Association.
- Eatwell, J., & Wilson, I. (2007). The effective use of psychometric assessments in decision making. In I. M. Evans, J. J. Rucklidge, & M. O'Driscoll (Eds.), *Professional practice of psychology in Aotearoa New Zealand* (pp. 317-354). Wellington, NZ: New Zealand Psychological Society.
- Ellsworth, J. R., Lambert, M. J., & Johnson, J. (2006). A comparison of the Outcome Questionnaire 45 and Outcome Questionnaire 30 in classification and prediction of treatment outcome. *Clinical Psychology & Psychotherapy*, *13*(6), 380-391.
- Evans, I. M. (1986). Response structure and the triple response mode concept in behavioral assessment. In R. O. Nelson, & S. C. Hayes (Eds.), *Conceptual foundations of behavioral assessment* (pp. 131-155). New York, NY: Guilford.
- Evans, I. M. (1999). Child-focused behavioral assessment and modification. *Journal of Clinical Child Psychology*, *28*(4), 493-501.

- Evans, I. M. (2002). Clinical psychology in early 21st century Aotearoa/New Zealand: Introduction to the special issue. *New Zealand Journal of Psychology, 31*(2), 50-53.
- Evans, I. M. (2005). *Self-assessment of therapeutic progress and change*. Unpublished Manuscript. School of Psychology, Massey University. Palmerston North, NZ.
- Evans, I. M. (2013). *How and why people change: Foundations of psychological therapy*. New York, NY: Oxford University.
- Evans, I. M., & Fitzgerald, J. (2007). Integrating research and practice in professional psychology: Models and paradigms. In I. M. Evans, J. J. Rucklidge, & M. O'Driscoll (Eds.), *Professional practice of psychology in Aotearoa New Zealand*. Wellington, NZ: New Zealand Psychological Society.
- Evans, I. M., Meyer, L. M., Kurkjian, J. A., & Kishi, G. S. (1998). An evaluation of behavioural interrelationships in child behaviour therapy In J. C. Witt, S. N. Elliott, & F. N. Gresham (Eds.), *Handbook of behavior therapy in education* (pp. 189-215). New York, NY: Plenum.
- Eysenck, H. J. (1952). The effects of psychotherapy: An evaluation. *Journal of Consulting Psychology, 16*(5), 319 - 324.
- Fennell, M. J., & Teasdale, J. D. (1987). Cognitive therapy for depression: Individual differences and the process of change. *Cognitive Therapy and Research, 11*(2), 253-271.

- Field, A. P. (2009). *Discovering statistics using SPSS : (and sex and drugs and rock 'n' roll)* (3rd ed.). London, UK: Sage.
- Field, A. P. (2013). *Discovering statistics using IBM SPSS statistics : and sex and drugs and rock 'n' roll* (4th ed.). London, UK: Sage.
- Field, S. D., Barkham, M., Shapiro, D. A., & Stiles, W. B. (1994). Assessment of assimilation in psychotherapy: A quantitative case study of problematic experiences with a significant other. *Journal of Counseling Psychology, 41*(3), 397-397.
- Finch, A. E., Lambert, M. J., & Schaalje, B. G. (2001). Psychotherapy quality control: The statistical generation of expected recovery curves for integration into an early warning system. *Clinical Psychology & Psychotherapy, 8*(4), 231-242.
- Fitzgerald, J. (2007). Measuring outcomes. In I. M. Evans, J. J. Rucklidge, & M. O'Driscoll (Eds.), *Professional practice of psychology in Aotearoa/New Zealand*. Wellington, NZ: New Zealand Psychological Society.
- Fletcher, A. (2011). *The expectations of experienced and novice clinical psychologists regarding course of change for clients undertaking successful cognitive behavioural psychotherapy*. (Unpublished Master's Thesis). Massey University. Palmerston North, NZ.
- Friedman, E. S., & Thase, M. E. (2007). Depression-focused psychotherapies. In G. O. Gabbard (Ed.), *Gabbard's treatments of psychiatric disorders* (4th ed., pp. 409-431). Arlington, VA: American Psychiatric Association.

- Froyd, J. E. (1996). A review of practices of psychotherapy outcome measurement. *Journal of Mental Health, 5*(1), 11-16.
- Gaines, P., Bower, A., Buckingham, W., Eagar, K., Burgess, P., & Green, J. (2003). *New Zealand mental health classification and outcomes study: Final report*. Auckland, NZ: Health Research Council of New Zealand.
- Garland, A. F., Kruse, M., & Aarons, G. A. (2003). Clinicians and outcome measurement: What's the use? *The Journal of Behavioral Health Services and Research, 30*(4), 393-405.
- Gaynor, S. T., Weersing, V. R., Kolko, D. J., Birmaher, B., Heo, J., & Brent, D. A. (2003). The prevalence and impact of large sudden improvements during adolescent therapy for depression: A comparison across cognitive-behavioral, family, and supportive therapy. *Journal of Consulting and Clinical Psychology, 71*(2), 386-393.
- George, D., & Mallery, P. (Eds.). (2014). *IBM SPSS statistics 21 step by step: A simple guide and reference* (13th ed.). Boston, NJ: Pearson.
- Gianakis, M., & Carey, T. A. (2011). An interview study investigating experiences of psychological change without psychotherapy. *Psychology and Psychotherapy: Theory, Research and Practice, 84*(4), 442-457.
- Goldfried, M. R. (2011). Generating research questions from clinical experience: Therapists' experiences in using CBT for panic disorder. *The Behavior Therapist, 34*(4), 57 - 62.

- Goldfried, M. R., & Wolfe, B. E. (1996). Psychotherapy practice and research: Repairing a strained alliance. *American Psychologist, 51*(10), 1007-1016.
- Goodman, J. D., McKay, J. R., & DePhilippis, D. (2013). Progress monitoring in mental health and addiction treatment: A means of improving care. *Professional Psychology: Research and Practice, 44*(4), 231 - 246.
- Grafanaki, S., & McLeod, J. (1999). Narrative processes in the construction of helpful and hindering events in experiential psychotherapy. *Psychotherapy Research, 9*(3), 289-303.
- Graham, S. E., Singer, J. D., & Willett, J. B. (2009). Modeling individual change over time. In A. Maydeu-Olivares, & R. E. Millsap (Eds.), *The SAGE Handbook of Quantitative Methodology for the Social Sciences*. London, UK: Sage.
- Greenberg, L. S. (2007). A guide to conducting a task analysis of psychotherapeutic change. *Psychotherapy Research, 17*(1), 15-30.
- Greenberg, R. P., Constantino, M. J., & Bruce, N. (2006). Are patient expectations still relevant for psychotherapy process and outcome? *Clinical Psychology Review, 26*(6), 657-678.
- Grosbeck, C. J. (1989). CG Jung and the shaman's vision. *Journal of Analytical Psychology, 34*(3), 255-275.
- Groth-Marnat, G. (Ed.). (2003). *Handbook of psychological assessment* (4th ed.). Hoboken, NJ: Wiley.

- Grundy, D. (2008). The presence of stigma among users of the MMORPG RMT: A hypothetical case approach. *Games and Culture*, 3(2), 225-247.
- Haas, E., Hill, R. D., Lambert, M. J., & Morrell, B. (2002). Do early responders to psychotherapy maintain treatment gains? *Journal of Clinical Psychology*, 58(9), 1157-1172.
- Hamilton, J. D., Kendall, P. C., Gosch, E., Furr, J. M., & Sood, E. (2008). Flexibility within fidelity. *Journal of the American Academy of Child & Adolescent Psychiatry*, 47(9), 987-993.
- Hannan, C., Lambert, M. J., Harmon, C., Nielsen, S. L., Smart, D. W., Shimokawa, K., & Sutton, S. W. (2005). A lab test and algorithms for identifying clients at risk for treatment failure. *Journal of Clinical Psychology*, 61(2), 155-163.
- Hansen, N. B., & Lambert, M. J. (2003). An evaluation of the dose–response relationship in naturalistic treatment settings using survival analysis. *Mental Health Services Research*, 5(1), 1-12.
- Hansen, N. B., Lambert, M. J., & Forman, E. M. (2002). The psychotherapy dose response effect and its implications for treatment delivery services. *Clinical Psychology: Science and Practice*, 9(3), 329-343.
- Hardy, G. E., Cahill, J., Stiles, W. B., Ispan, C., Macaskill, N., & Barkham, M. (2005). Sudden gains in cognitive therapy for depression: A replication and extension. *Journal of Consulting and Clinical Psychology*, 73(1), 59-67.

- Harmon, S. C., Lambert, M. J., Smart, D. M., Hawkins, E., Nielsen, S. L., Slade, K., & Lutz, W. (2007). Enhancing outcome for potential treatment failures: Therapist-client feedback and clinical support tools. *Psychotherapy Research, 17*(4), 379-392.
- Hatfield, D., McCullough, L., Frantz, S. H. B., & Krieger, K. (2010). Do we know when our clients get worse? An investigation of therapists' ability to detect negative client change. *Clinical Psychology & Psychotherapy, 17*(1), 25-32.
- Hatfield, D. R., & Ogles, B. M. (2004). The use of outcome measures by psychologists in clinical practice. *Professional Psychology: Research and Practice, 35*(5), 485-491.
- Hatfield, D. R., & Ogles, B. M. (2007). Why some clinicians use outcome measures and others do not. *Administration and Policy in Mental Health and Mental Health Services Research, 34*(3), 283-291.
- Hawkins, E. J., Lambert, M. J., Vermeersch, D. A., Slade, K. L., & Tuttle, K. C. (2004). The therapeutic effects of providing patient progress information to therapists and patients. *Psychotherapy Research, 14*(3), 308-327.
- Hayes, A. M., Feldman, G. C., Beevers, C. G., Laurenceau, J. P., Cardaciotto, L. A., & Lewis-Smith, J. (2007a). Discontinuities and cognitive changes in an exposure-based cognitive therapy for depression. *Journal of Consulting and Clinical Psychology, 75*(3), 409-421.

- Hayes, A. M., Laurenceau, J. P., Feldman, G., Strauss, J. L., & Cardaciotto, L. A. (2007b). Change is not always linear: The study of nonlinear and discontinuous patterns of change in psychotherapy. *Clinical Psychology Review, 27*(6), 715-723.
- Hayes, S. A., Miller, N. A., Hope, D. A., Heimberg, R. G., & Juster, H. R. (2008). Assessing client progress session by session in the treatment of social anxiety disorder: The social anxiety session change index. *Cognitive and Behavioral Practice, 15*(2), 203-211.
- Hayes, S. C., Strosahl, K. D., & Wilson, K. G. (1999). *Acceptance and commitment therapy: An experiential approach to behavior change*. New York, NY: Guilford.
- Health and Care Professions Council. (2012). *Standards of proficiency: Practitioner psychologists*. London, UK: Health and Care Professions Council.
- Heck, R. H., Thomas, S. L., & Tabata, L. N. (2010). *Multilevel and longitudinal modeling with IBM SPSS*. New York: Routledge.
- Hedeker, D. R. (2004). An introduction to growth modeling In D. Kaplan (Ed.), *The Sage handbook of quantitative methodology for the social sciences* (pp. 216 - 236). Thousand Oaks, CA: Sage.
- Heimberg, R. G., & Becker, R. E. (2002). *Cognitive-behavioral group therapy for social phobia: Basic mechanisms and clinical strategies*. New York, NY: Guilford.

- Hofmann, S. G., Schulz, S. M., Meuret, A. E., Moscovitch, D. A., & Suvak, M. (2006). Sudden gains during therapy of social phobia. *Journal of Consulting and Clinical Psychology, 74*(4), 687-697.
- Hollon, S. D., & Beck, A. T. (2013). Cognitive and cognitive-behavioral therapies. In M. J. Lambert (Ed.), *Bergin and Garfield's handbook of psychotherapy and behavioral change*. Hoboken, NJ: Wiley.
- Honos-Webb, L., Stiles, W., Greenberg, L., & Goldman, R. (1998). Assimilation analysis of process-experiential psychotherapy: A comparison of two cases. *Psychotherapy Research, 8*(3), 264-286.
- Hopko, D. R., Armento, M. E. A., Cantu, M. S., Chambers, L. L., & Lejuez, C. (2003). The use of daily diaries to assess the relations among mood state, overt behavior, and reward value of activities. *Behaviour Research and Therapy, 41*(10), 1137-1148.
- Howard, K. I., Kopta, S., Krause, M. S., & Orlinsky, D. E. (1986). The dose-effect relationship in psychotherapy. *American Psychologist, 41*(2), 159-164.
- Howard, K. I., Lueger, R. J., Maling, M. S., & Martinovich, Z. (1993). A phase model of psychotherapy outcome: Causal mediation of change. *Journal of Consulting and Clinical Psychology, 61*(4), 678-685.
- Howard, K. I., Moras, K., Brill, P. L., Martinovich, Z., & Lutz, W. (1996). Evaluation of psychotherapy: Efficacy, effectiveness, and patient progress. *American Psychologist, 51*(10), 1059-1064.

Humphreys, C. L., Rubin, J. S., Knudson, R. M., & Stiles, W. B. (2005). The assimilation of anger in a case of dissociative identity disorder. *Counselling Psychology Quarterly*, *18*(2), 121-132.

IBM Corp. (Released 2012). IBM SPSS statistics for Windows, Version 21.0. Armonk, NY: IBM Corp.

IBM SPSS Analytic Catalyst. (n.d). Adjusted McFadden's R-squared change. Retrieved May 13, 2014, from http://pic.dhe.ibm.com/infocenter/spssas/v1r0m0/index.jsp?topic=%2Fcom.ibm.spss.analyticcatalyst.help%2Fanalytic_catalyst%2Fmcfadden_r_squared.html

Ilardi, S. S., & Craighead, W. E. (1994). The role of nonspecific factors in cognitive behavior therapy for depression. *Clinical Psychology: Science and Practice*, *1*(2), 138-155.

Ingram, R. E., Overbey, T., & Fortier, M. (2001). Individual differences in dysfunctional automatic thinking and parental bonding: Specificity of maternal care. *Personality and Individual Differences*, *30*(3), 401-412.

Jacobson, N., & Truax, P. (1991). Clinical significance: A statistical approach to defining meaningful change in psychotherapy research. *Journal of Consulting and Clinical Psychology*, *59*(1), 12-19.

Jacobson, N. S., Follette, W. C., & Revenstorf, D. (1984). Psychotherapy outcome research: Methods for reporting variability and evaluating clinical significance. *Behavior Therapy*, *15*(4), 336-352.

- Jacobson, N. S., Roberts, L. J., Berns, S. B., & McGlinchey, J. B. (1999). Methods for defining and determining the clinical significance of treatment effects: Description, application, and alternatives. *Journal of Consulting and Clinical Psychology, 67*(3), 300-307.
- Johnson, J., Lambert, M., & Sullivan, E. (2001). *Early response and psychotherapy outcomes*. Paper presented at the 109th Annual Meeting of the American Psychological Association, San Francisco, CA.
- Johnson, R. W. (1965). Number of interviews, diagnosis and success of counseling. *Journal of Counseling Psychology, 12*(3), 248-251.
- Jung, C. G. (1968). *Collected works of C. G. Jung* (2nd ed., Vol. 12, pp. 1–39). Princeton, NJ: Princeton University.
- Kadera, S. W., Lambert, M. J., & Andrews, A. A. (1996). How much therapy is really enough?: A session-by-session analysis of the psychotherapy dose-effect relationship. *Journal of Psychotherapy Practice and Research, 5*(2), 132-151.
- Kahn, J. H., & Schneider, W. J. (2013). It's the Destination and It's the Journey: Using Multilevel Modeling to Assess Patterns of Change in Psychotherapy. *Journal of Clinical Psychology, 69*(6), 543-570.
- Kazantzis, N., & Deane, F. P. (1998). Theoretical orientations of New Zealand psychologists: An international comparison. *Journal of Psychotherapy Integration, 8*(2), 97-113.

- Kazantzis, N., Whittington, C., & Dattilio, F. (2010). Meta-analysis of homework effects in cognitive and behavioral therapy: A replication and extension. *Clinical Psychology: Science and Practice, 17*(2), 144-156.
- Kazdin, A. E. (1999). The meanings and measurement of clinical significance. *Journal of Consulting and Clinical Psychology, 67*(3), 332-339.
- Kazdin, A. E. (2007). Mediators and mechanisms of change in psychotherapy research. *Annual Review of Clinical Psychology, 3*, 1-27.
- Kazdin, A. E. (2009). Understanding how and why psychotherapy leads to change. *Psychotherapy Research, 19*(4-5), 418-428.
- Kelly, M. A. R., Roberts, J. E., & Bottonari, K. A. (2007). Non-treatment-related sudden gains in depression: The role of self-evaluation. *Behaviour Research and Therapy, 45*(4), 737-747.
- Kelly, M. A. R., Roberts, J. E., & Ciesla, J. A. (2005). Sudden gains in cognitive behavioral treatment for depression: when do they occur and do they matter? *Behaviour Research and Therapy, 43*(6), 703-714.
- Kelso, J. A. S. (1995). *Dynamic patterns: The self-organization of brain and behavior*. Cambridge, MA: The MIT Press.
- Kopta, S. M. (2003). The dose-effect relationship in psychotherapy: A defining achievement for Dr. Kenneth Howard. *Journal of Clinical Psychology, 59*(7), 727-733.

- Kopta, S. M., Howard, K. I., Lowry, J. L., & Beutler, L. E. (1994). Patterns of symptomatic recovery in psychotherapy. *Journal of Consulting and Clinical Psychology, 62*(5), 1009-1016.
- Kopta, S. M., Lueger, R. J., Saunders, S. M., & Howard, K. I. (1999). Individual psychotherapy outcome and process research: Challenges leading to greater turmoil or a positive transition? *Annual Review of Psychology, 50*(1), 441-469.
- Krause, M. S., Howard, K. I., & Lutz, W. (1998). Exploring individual change. *Journal of Consulting and Clinical Psychology, 66*(5), 838-845.
- Kuyken, W., Watkins, E., & Beck, A. T. (2007). Cognitive-behavioral therapy for mood disorders. In G. O. Gabbard, J. S. Beck, & J. Holmes (Eds.), *Oxford textbook of psychotherapy* (pp. 111-126). New York, NY: Oxford University.
- Lambert, M. J. (2005). Early response in psychotherapy: further evidence for the importance of common factors rather than “placebo effects”. *Journal of Clinical Psychology, 61*(7), 855-869.
- Lambert, M. J. (2010). *Prevention of treatment failure: The use of measuring, monitoring, and feedback in clinical practice*. Washington, DC: American Psychological Association.
- Lambert, M. J., & Finch, A. E. (1999). The Outcome Questionnaire. In M. E. Maruish (Ed.), *The use of psychological testing for treatment planning and outcome assessment* (2nd ed., pp. 831 - 870). Mahwah, NJ: Earlbaum.

- Lambert, M. J., Hansen, N. B., & Finch, A. E. (2001a). Patient-focused research: Using patient outcome data to enhance treatment effects. *Journal of Consulting and Clinical Psychology, 69*(2), 159-172.
- Lambert, M. J., & Hill, C. E. (1994). Assessing psychotherapy outcomes and processes. In A. E. Bergin, & S. L. Garfield (Eds.), *Handbook of psychotherapy and behavior change* (4 ed., pp. 72 - 113). Hoboken, NY: Wiley.
- Lambert, M. J., & Shimokawa, K. (2011). Collecting client feedback. *Psychotherapy, 48*(1), 72-79.
- Lambert, M. J., Whipple, J. L., Bishop, M. J., Vermeersch, D. A., Gray, G. V., & Finch, A. E. (2002a). Comparison of empirically derived and rationally derived methods for identifying patients at risk for treatment failure. *Clinical Psychology and Psychotherapy, 9*(3), 149-164.
- Lambert, M. J., Whipple, J. L., Hawkins, E. J., Vermeersch, D. A., Nielsen, S. L., & Smart, D. W. (2003). Is it time for clinicians to routinely track patient outcome? A meta-analysis. *Clinical Psychology: Science and Practice, 10*(3), 288-301.
- Lambert, M. J., Whipple, J. L., Smart, D. W., Vermeersch, D. A., & Nielsen, S. L. (2001b). The effects of providing therapists with feedback on patient progress during psychotherapy: Are outcomes enhanced? *Psychotherapy Research, 11*(1), 49-68.
- Lambert, M. J., Whipple, J. L., Vermeersch, D. A., Smart, D. W., Hawkins, E. J., Nielsen, S. L., & Goates, M. (2002b). Enhancing psychotherapy outcomes via providing

- feedback on client progress: A replication. *Clinical Psychology and Psychotherapy*, 9(2), 91-103.
- Laurenceau, J. P., Hayes, A. M., & Feldman, G. C. (2007). Statistical and methodological issues in the study of change in psychotherapy. *Clinical Psychology Review*, 27(6), 715–723.
- Lejuez, C., Hopko, D. R., LePage, J. P., Hopko, S. D., & McNeil, D. W. (2001). A brief behavioral activation treatment for depression. *Cognitive and Behavioral Practice*, 8(2), 164-175.
- Leon, S. C., Kopta, S. M., Howard, K. I., & Lutz, W. (1999). Predicting patients' responses to psychotherapy: Are some more predictable than others? *Journal of Consulting and Clinical Psychology*, 67(5), 698-704.
- Lincoln, Y., & Guba, E. (2002). Judging the quality of case study reports. In M. Huberman, & M. Miles (Eds.), *The qualitative researchers' companion* (pp. 205-216). Thousand Oaks, CA: Sage.
- Linley, P. A., & Joseph, S. (2004). Positive change following trauma and adversity: A review. *Journal of Traumatic Stress*, 17(1), 11-21.
- Lueger, R. J., Howard, K. I., Martinovich, Z., Lutz, W., Anderson, E. E., & Grissom, G. (2001). Assessing treatment progress of individual patients using expected treatment response models. *Journal of Consulting and Clinical Psychology*, 69(2), 150-158.

- Lueger, R. J., Lutz, W., & Howard, K. I. (2000). The predicted and observed course of psychotherapy for anxiety and mood disorders. *The Journal of Nervous and Mental Disease, 188*(3), 127 - 134.
- Lutz, W. (2003). Efficacy, effectiveness, and expected treatment response in psychotherapy. *Journal of Clinical Psychology, 59*(7), 745-750.
- Lutz, W., Lambert, M. J., Harmon, S. C., Tschitsaz, A., Schürch, E., & Stulz, N. (2006a). The probability of treatment success, failure and duration—what can be learned from empirical data to support decision making in clinical practice? *Clinical Psychology and Psychotherapy, 13*(4), 223-232.
- Lutz, W., Lowry, J., Kopta, S. M., Einstein, D. A., & Howard, K. I. (2001). Prediction of dose–response relations based on patient characteristics. *Journal of Clinical Psychology, 57*(7), 889-900.
- Lutz, W., Martinovich, Z., & Howard, K. I. (1999). Patient profiling: An application of random coefficient regression models to depicting the response of a patient to outpatient psychotherapy. *Journal of Consulting and Clinical Psychology, 67*(4), 571-577.
- Lutz, W., Saunders, S. M., Leon, S. C., Martinovich, Z., Kosfelder, J., Schulte, D., . . . Tholen, S. (2006b). Empirically and clinically useful decision making in psychotherapy: Differential predictions with treatment response models. *Psychological Assessment, 18*(2), 133-141.

- Lutz, W., Stulz, N., & Köck, K. (2009). Patterns of early change and their relationship to outcome and follow-up among patients with major depressive disorders. *Journal of Affective Disorders, 118*(1-3), 60-68.
- Maling, M. S., Gurtman, M. B., & Howard, K. I. (1995). The response of interpersonal problems to varying doses of psychotherapy. *Psychotherapy Research, 5*(1), 63-75.
- McNeilly, C. L., & Howard, K. I. (1991). The effects of psychotherapy: A reevaluation based on dosage. *Psychotherapy Research, 1*(1), 74-78.
- Meites, T. M., Ingram, R. E., & Siegle, G. J. (2012). Unique and shared aspects of affective symptomatology: The role of parental bonding in depression and anxiety symptom profiles. *Cognitive Therapy and Research, 36*(3), 173 - 181.
- Mellor-Clark, J., Barkham, M., Evans, C., Margison, F., McGrath, G., & Connell, J. (1997). *Talking effectiveness: Providers' and purchasers' perspectives on outcomes*. Leeds, UK: Psychological Therapies Research Centre, University of Leeds.
- Mellsop, G., & Smith, M. (2010). Outcome measures in New Zealand. In T. Trauer (Ed.), *Outcome measurement in mental health* (pp. 26-31). Cambridge, UK: Cambridge University.
- Meyer, T. J., Miller, M. L., Metzger, R. L., & Borkovec, T. D. (1990). Development and validation of the Penn State Worry Questionnaire. *Behaviour Research and Therapy, 28*(6), 487-495.

- Michie, S., & Johnston, M. (2012). Theories and techniques of behaviour change: Developing a cumulative science of behaviour change. *Health Psychology Review*, 6(1), 1-6.
- Miller, R. C., & Berman, J. S. (1983). The efficacy of cognitive behavior therapies: A quantitative review of the research evidence. *Psychological Bulletin*, 94(1), 39-53.
- Miller, S. D., Duncan, B. L., & Hubble, M. A. (2004). Beyond integration: The triumph of outcome over process in clinical practice. *Psychotherapy in Australia*, 10(2), 2-19.
- Miller, S. D., Duncan, B. L., Sorrell, R., & Brown, G. S. (2005). The partners for change outcome management system. *Journal of Clinical Psychology*, 61(2), 199-208.
- Mohr, D. C. (1995). Negative outcome in psychotherapy: A critical review. *Clinical Psychology: Science and Practice*, 2(1), 1-27.
- Mohr, D. C., Beutler, L. E., Engle, D., Shoham-Salomon, V., Bergan, J., Kaszniak, A. W., & Yost, E. B. (1990). Identification of patients at risk for nonresponse and negative outcome in psychotherapy. *Journal of Consulting and Clinical Psychology*, 58(5), 622-628.
- Nathan, P. E., Stuart, S. P., & Dolan, S. L. (2000). Research on psychotherapy efficacy and effectiveness: Between Scylla and Charybdis? *Psychological Bulletin*, 126(6), 964-981.
- New Zealand Psychologists Board. (2011). *Core competencies: For the practice of psychology in New Zealand*. Wellington, NZ: New Zealand Psychologists Board.

- Newman, F. L., Ciarlo, J. A., & Carpenter, D. (1999). Guidelines for selecting psychological instruments for treatment planning and outcome assessment. In M. E. Maruish (Ed.), *The use of psychological testing for treatment planning and outcomes assessment* (2nd ed., pp. 153-170). Mahwah, NJ: Lawrence Erlbaum Associates.
- Newnham, E. A., Hooke, G. R., & Page, A. C. (2010). Monitoring treatment response and outcomes using the World Health Organization's Wellbeing Index in psychiatric care. *Journal of Affective Disorders, 122*(1-2), 133-138.
- Newnham, E. A., & Page, A. C. (2007). Client-focused research: New directions in outcome assessment. *Behaviour Change, 24*(1), 1-6.
- Newnham, E. A., & Page, A. C. (2010). Bridging the gap between best evidence and best practice in mental health. *Clinical Psychology Review, 30*(1), 127-142.
- Norcross, J. C., Krebs, P. M., & Prochaska, J. O. (2011). Stages of change. *Journal of Clinical Psychology: In Session, 67*(2), 1-12.
- Oei, T. P. S., Bullbeck, K., & Campbell, J. M. (2006). Cognitive change process during group cognitive behaviour therapy for depression. *Journal of Affective Disorders, 92*(2-3), 231-241.
- Orlinsky, D. E., Ronnestad, M. H., & Willutzki, U. (2004). Fifty years of psychotherapy process-outcome research: Continuity and change. In M. J. Lambert (Ed.), *Bergin and Garfield's handbook of psychotherapy and behavior change* (5 ed., pp. 307-389). New York, NY: Wiley.

- Osatuke, K., Humphreys, C. L., Glick, M. J., Graff-Reed, R. L., Mack, L. T. M. K., & Stiles, W. B. (2005). Vocal manifestations of internal multiplicity: Mary's voices. *Psychology and Psychotherapy: Theory, Research and Practice, 78*(1), 21-44.
- Parabiaghi, A., Barbato, A., D'avanzo, B., Erlicher, A., & Lora, A. (2005). Assessing reliable and clinically significant change on Health of the Nation Outcome Scales: method for displaying longitudinal data. *Australian and New Zealand Journal of Psychiatry, 39*(8), 719-724.
- Patchett-Andersen, L. S. (1997). *Clinical psychologists' opinions about and uses of tests, assessment, and clinical intervention application*. (Unpublished Master's Thesis). Massey University. Palmerston North, NZ.
- Patterson, A. (1999). The publication of case studies and confidentiality—an ethical predicament. *Psychiatric Bulletin, 23*(9), 562-564.
- Paulson, B. L., Truscott, D., & Stuart, J. (1999). Clients' perceptions of helpful experiences in counseling. *Journal of Counseling Psychology, 46*(3), 317-324.
- Phelps, R., Eisman, E. J., & Kohout, J. (1998). Psychological practice and managed care: Results of the CAPP practitioner survey. *Professional Psychology: Research and Practice, 29*(1), 31-36.
- Present, J., Crits-Christoph, P., Connolly Gibbons, M. B., Hearon, B., Ring-Kurtz, S., Worley, M., & Gallop, R. (2008). Sudden gains in the treatment of generalized anxiety disorder. *Journal of Clinical Psychology, 64*(1), 119-126.

- Prochaska, J. O., & DiClemente, C. C. (1982). Transtheoretical therapy: Toward a more integrative model of change. *Psychotherapy: Theory, Research and Practice, 19*(3), 276-288.
- Prochaska, J. O., & DiClemente, C. C. (1983). Stages and processes of self-change of smoking: toward an integrative model of change. *Journal of Consulting and Clinical Psychology, 51*(3), 390-395.
- Quality Assurance Project. (1982). A treatment outline for agoraphobia. *Australian and New Zealand Journal of Psychiatry, 16*(1), 25-33.
- Quality Assurance Project. (1983). A treatment outline for depressive disorders. *Australian and New Zealand Journal of Psychiatry, 17*, 129-146.
- Quality Assurance Project. (1984). A treatment outline for the management of schizophrenia. *Australian and New Zealand Journal of Psychiatry, 18*(1), 19-38.
- Quality Assurance Project. (1985a). Treatment outlines for the management of anxiety states. *Australian and New Zealand Journal of Psychiatry, 19*(2), 138-151.
- Quality Assurance Project. (1985b). Treatment outlines for the management of obsessive-compulsive disorders. *Australian and New Zealand Journal of Psychiatry, 19*(3), 240-253.
- Raftery, A. E. (1993). Bayesian model selection in structural equation models. In K. A. Bollen, & J. S. Long (Eds.), *Testing Structural Equation Models*. Beverly Hills, CA: Sage.

- Rogers, C. R. (1958). A process conception of psychotherapy. *American Psychologist, 13*(4), 142-149.
- Rosen, G. M., & Davison, G. C. (2003). Psychology should list empirically supported principles of change (ESPs) and not credential trademarked therapies or other treatment packages. *Behavior Modification, 27*(3), 300-312.
- Ryan, R. M., Lynch, M. F., Vansteenkiste, M., & Deci, E. L. (2011). Motivation and autonomy in counseling, psychotherapy, and behavior change: A look at theory and practice. *The Counseling Psychologist, 39*(2), 193-260.
- Samoilov, A., & Goldfried, M. R. (2000). Role of emotion in cognitive behavior therapy. *Clinical Psychology: Science and Practice, 7*(4), 373-385.
- Sander, L. (1988). The event-structure of regulation in the neonate-caregiver system as a biological background for early organization of psychic structure. In A. Goldberg (Ed.), *Frontiers in self-psychology* (pp. 64-77). Hillsdale, NJ: Analytic Press.
- Sander, L. (1994). Paradox and resolution: From the beginning. In J. Noshpitz, S. Greenspan, S. Weider, & J. Osofsky (Eds.), *Handbook of child and adolescent psychiatry* (Vol. 1, pp. 152-159). New York, NY: Wiley.
- Sapyta, J., Riemer, M., & Bickman, L. (2005). Feedback to clinicians: Theory, research, and practice. *Journal of Clinical Psychology, 61*(2), 145-153.
- Schachter, S., & Singer, J. (1962). Cognitive, social, and physiological determinants of emotional state. *Psychological Review, 69*(5), 379.

- Schneier, F. R., Mellman, L. A., & Spiegel, D. (2007). Anxiety disorders, dissociative disorders and adjustment disorders. In G. O. Gabbard (Ed.), *Gabbard's treatments of psychiatric disorders* (4 ed., Vol. 477-478). Arlington, VA: American Psychiatric Publishing.
- Seeman, J. (1954). Counselor judgments of therapeutic process and outcome:. In C. R. Rogers, & R. F. Dymond (Eds.), *Psychotherapy and personality change* (pp. 99-108). Chicago, IL: University of Chicago.
- Seligman, M. E. P. (1995). The effectiveness of psychotherapy: The Consumer Reports study. *American Psychologist*, *50*(12), 965-974.
- Shek, D. T., & Ma, C. (2011). Longitudinal data analyses using linear mixed models in SPSS: Concepts, procedures and illustrations. *The Scientific World Journal*, *11*, 42-76.
- Simons, A. D., Gordon, J. S., Monroe, S. M., & Thase, M. E. (1995). Toward an integration of psychologic, social, and biologic factors in depression: Effects on outcome and course of cognitive therapy. *Journal of Consulting and Clinical Psychology*, *63*(3), 369-377.
- Singer, J. D., & Willett, J. B. (2003). *Applied longitudinal data analysis : modeling change and event occurrence*. Oxford, UK: Oxford University.
- Smith, M. L., Glass, G. V., & Miller, T. I. (1980). *The benefits of psychotherapy*. Baltimore, MD: Johns Hopkins University.

- Society for Psychotherapy Research. (1993). Review of contents assessed before and after treatment and their measures. *Obtained through Barkham, M., Evans, C., Margison, F., McGrath, G, Mellor-Clark, J., Milne, D., & Connell, J. (1998). The rationale for developing and implementing core outcome batteries for routine use in service. Journal of Mental Health, 7(1), 35-47.*
- Speer, D. (1992). Clinically significant change: Jacobson and Truax (1991) revisited. *Journal of Consulting and Clinical Psychology, 60(3), 402-408.*
- Sperry, L., & Pies, R. (2010). Writing about clients: Ethical considerations and options. *Counseling and Values, 54(2), 88-102.*
- Spielmann, G. I., Masters, K. S., & Lambert, M. J. (2006). A comparison of rational versus empirical methods in the prediction of psychotherapy outcome. *Clinical Psychology and Psychotherapy, 13(3), 202-214.*
- Staats, A. W. (1968). *Learning, language and cognition*. New York, NY: Holt, Rinehart & Winston.
- Staats, A. W. (1975). *Social behaviorism*. Homewood, IL: Dorsey.
- Startup, H. M., & Erickson, T. M. (2006). The Penn State worry questionnaire (PSWQ). In G. C. L. Davey, & A. Wells (Eds.), *Worry and its psychological disorders: Theory, assessment and treatment* (pp. 101–119). West Sussex, UK: Wiley.
- Steenbarger, B. N. (1992). Toward science-practice integration in brief counseling and therapy. *The Counseling Psychologist, 20(3), 403-450.*

- Steenbarger, B. N. (1994). Duration and outcome in psychotherapy: An integrative review. *Professional Psychology: Research and Practice, 25*(2), 111-119.
- Stern, D. N., Bruschiweiler-Stern, N., Harrison, A. M., Lyons-Ruth, K., Morgan, A. C., Nahum, J. P., . . . Tronick, E. Z. (1998). The process of therapeutic change involving implicit knowledge: Some implications of developmental observations for adult psychotherapy. *Infant Mental Health Journal, 19*(3), 300-308.
- Stewart, J. W., Quitkin, F. M., McGrath, P. J., Amsterdam, J., Fava, M., Fawcett, J., . . . Roback, P. (1998). Use of pattern analysis to predict differential relapse of remitted patients with major depression during 1 year of treatment with fluoxetine or placebo. *Archives of General Psychiatry, 55*(4), 334-343.
- Stiles, W., Meshot, C., Anderson, T., & Sloan, W. (1992). Assimilation of problematic experiences: The case of John Jones. *Psychotherapy Research, 2*(2), 81-101.
- Stiles, W. B. (2001). Assimilation of problematic experiences. *Psychotherapy: Theory, Research, Practice, Training, 38*(4), 462-465.
- Stiles, W. B. (2006). Assimilation and the process of outcome: Introduction to a special section. *Psychotherapy Research, 16*(4), 389-392.
- Stiles, W. B., Elliott, R., Llewelyn, S. P., Firth-Cozens, J. A., Margison, F. R., Shapiro, D. A., & Hardy, G. (1990). Assimilation of problematic experiences by clients in psychotherapy. *Psychotherapy: Theory, Research, Practice, Training, 27*(3), 411-420.

- Stiles, W. B., Honos-Webb, L., & Surko, M. (1998). Responsiveness in psychotherapy. *Clinical Psychology: Science and Practice, 5*(4), 439-458.
- Stiles, W. B., Leach, C., Barkham, M., Lucock, M., Iveson, S., Shapiro, D. A., . . . Hardy, G. E. (2003). Early sudden gains in psychotherapy under routine clinic conditions: Practice-based evidence. *Journal of Consulting and Clinical Psychology, 71*(1), 14-21.
- Stiles, W. B., Leiman, M., Shapiro, D. A., Hardy, G. E., Barkham, M., Detert, N. B., & Llewelyn, S. P. (2006). What does the first exchange tell? Dialogical sequence analysis and assimilation in very brief therapy. *Psychotherapy Research, 16*(4), 408-421.
- Stiles, W. B., Morrison, L. A., Haw, S. K., Harper, H., Shapiro, D. A., & Firth-Cozens, J. (1991). Longitudinal study of assimilation in exploratory psychotherapy. *Psychotherapy: Theory, Research, Practice, Training, 28*(2), 195-206.
- Strassberg, D. S., Anchor, K. N., Cunningham, J., & Elkins, D. (1977). Successful outcome and number of sessions: When do counselors think enough is enough? *Journal of Counseling Psychology, 24*(6), 477-480.
- Stulz, N., Lutz, W., Leach, C., Lucock, M., & Barkham, M. (2007). Shapes of early change in psychotherapy under routine outpatient conditions. *Journal of Consulting and Clinical Psychology, 75*(6), 864 - 874.

- Tang, T. Z., & DeRubeis, R. J. (1999a). Reconsidering rapid early response in cognitive behavioral therapy for depression. *Clinical Psychology: Science and Practice*, 6(3), 283-288.
- Tang, T. Z., & DeRubeis, R. J. (1999b). Sudden gains and critical sessions in cognitive-behavioral therapy for depression. *Journal of Consulting and Clinical Psychology*, 67(6), 894-904.
- Tang, T. Z., DeRubeis, R. J., Beberman, R., & Pham, T. (2005). Cognitive changes, critical sessions, and sudden gains in cognitive-behavioral therapy for depression. *Journal of Consulting and Clinical Psychology*, 73(1), 168-172.
- Tang, T. Z., Luborsky, L., & Andrusyna, T. (2002). Sudden gains in recovering from depression: Are they also found in psychotherapies other than cognitive-behavioral therapy? *Journal of Consulting and Clinical Psychology*, 70(2), 444-447.
- Teachman, B. A., Drabick, D. A., Hershenberg, R., Vivian, D., Wolfe, B. E., & Goldfried, M. R. (2012). Bridging the gap between clinical research and clinical practice: Introduction to the special section. *Psychotherapy*, 49(2), 97.
- Thompson, M., Thompson, L., & Gallagher-Thompson, D. (1995). Linear and nonlinear changes in mood between psychotherapy sessions: Implications for treatment outcome and relapse risk. *Psychotherapy Research*, 5(4), 327-336.
- Timulak, L. (2008). *Research in psychotherapy and counselling*. London, UK: SAGE.

- Timulak, L., & Elliott, R. (2003). Empowerment events in process--experiential psychotherapy of depression: An exploratory qualitative analysis. *Psychotherapy Research, 13*(4), 443-460.
- Trauer, T. (2010a). Assessment of change in outcome measurement. In T. Trauer (Ed.), *Outcome measurement in mental health: Theory and practice* (pp. 206-218). Cambridge, UK: Cambridge University.
- Trauer, T. (2010b). Outcome measurement in adult mental health services. In T. Trauer (Ed.), *Outcome measurement in mental health: Theory and practice* (pp. 116-125). Cambridge, UK: Cambridge University.
- Trauer, T. (2010c). *Outcome measurement in mental health: theory and practice*. Cambridge, UK: Cambridge University.
- Tronick, E. Z., Bruschiweiler-Stern, N., Harrison, A. M., Lyons-Ruth, K., Morgan, A. C., Nahum, J. P., . . . Stern, D. N. (1998). Dyadically expanded states of consciousness and the process of therapeutic change. *Infant Mental Health Journal, 19*(3), 290-299.
- Turk, C. L., & Wolanin, A. T. (2006). Assessment of generalized anxiety disorder. In G. C. L. Davey, & A. Wells (Eds.), *Worry and it's psychological disorders: theory, assessment, and treatment*. West Sussex, England: Wiley.
- Uhlenhuth, E. H., & Duncan, D. B. (1968). Subjective change with medical student therapists: II. Some determinants of change in psychoneurotic outpatients. *Archives of General Psychiatry, 18*(5), 532-540.

- Vallacher, R. R., Read, S. J., & Nowak, A. (2002). The dynamical perspective in personality and social psychology. *Personality and Social Psychology Review*, 6(4), 264-273.
- Vittengl, J. R., Clark, L. A., & Jarrett, R. B. (2005). Validity of sudden gains in acute phase treatment of depression. *Journal of Consulting and Clinical Psychology*, 73(1), 173-182.
- Voeltz, L. M., & Evans, I. M. (1982). The assessment of behavioral interrelationships in child behavior therapy. *Behavioral Assessment*, 4, 131-165.
- Watson, D., Clark, L. A., & Tellegen, A. (1988). Development and validation of brief measures of positive and negative affect: The PANAS scales. *Journal of Personality and Social Psychology*, 54(6), 1063-1070.
- Weisz, J. R., Donenberg, G. R., Han, S. S., & Weiss, B. (1995). Bridging the gap between laboratory and clinic in child and adolescent psychotherapy. *Journal of Consulting and Clinical Psychology*, 63(5), 688-701.
- Weitz, L. J., Abramowitz, S. I., Steger, J. A., Calabria, F. M., Conable, M., & Yarus, G. (1975). Number of sessions and client-judged outcome: The more the better? *Psychotherapy: Theory, Research and Practice*, 12(4), 337-340.
- Wells, A. (1997). *Cognitive therapy of anxiety disorders : a practice manual and conceptual guide*. Hoboken, NY: Wiley.

- Wells, A. (2006). Metacognitive therapy for worry and generalized anxiety disorder. In G. Davey, & A. Wells (Eds.), *Worry and its psychological disorders: theory, assessment and treatment* (pp. 179-200). Hoboken, NJ: Wiley.
- Westbrook, D., Kennerley, H., & Kirk, J. (Eds.). (2011). *An introduction to cognitive behaviour therapy* (Second ed.). London, UK: Sage.
- Westra, H. A., Dozois, D. J. A., & Marcus, M. (2007). Expectancy, homework compliance, and initial change in cognitive-behavioral therapy for anxiety. *Journal of Consulting and Clinical Psychology, 75*(3), 363-373.
- Whipple, J. L., & Lambert, M. J. (2011). Outcome measures for practice. *Annual Review of Clinical Psychology, 7*, 87-111.
- Whipple, J. L., Lambert, M. J., Vermeersch, D. A., Smart, D. W., Nielsen, S. L., & Hawkins, E. J. (2003). Improving the effects of psychotherapy: The use of early identification of treatment and problem-solving strategies in routine practice. *Journal of Counseling Psychology, 50*(1), 59-68.
- Wing, J., Beevor, A., Curtis, R., Park, S., Hadden, S., & Burns, A. (1998). Health of the Nation Outcome Scales (HoNOS). Research and development. *The British Journal of Psychiatry, 172*(1), 11-18.
- Wolfe, B. E. (2012). Healing the research–practice split: Let's start with me. *Psychotherapy, 49*(2), 101.

APPENDIX A: ELECTRONIC TASK

Information Sheet for Participants

Who is doing this research?

My name is Amber Fletcher and I am a student at Massey University. I am conducting this research project as a requirement for the completion of the Doctoral Programme in Clinical Psychology. The supervisor for this project is Professor Ian Evans.

What is this research about?

This project aims to find out what expectations experienced clinicians and students studying Clinical Psychology have in regards to client progress and outcomes during and following CBT treatment of particular disorders.

Who can participate?

Anyone interested in CBT and in clinical measurement. You do not have to be an experienced clinician as we are equally interested in the judgment of students and trainees. Your professional background is not important as long as you have some familiarity with standard CBT treatment approaches for clients with anxiety and depression. Actual level of experience, if any, is not important, but we will ask you a few demographic questions relevant to your background.

Your input is greatly appreciated. As a thank you for taking the time to complete our task, we are giving you the chance to win a \$100 [amazon.com](https://www.amazon.com) voucher. If you would like to be in the draw to win this voucher, please submit your email address in the provided area upon completion. Your email will be kept separate from your completed task, and is only used as a means to enter you in the draw.

Your rights as a participant:

You are under no obligation to accept this invitation to complete this task. Completion and submission of the questionnaire implies that you give your consent to participate. You are free to exit the task at any time. Information resulting from this questionnaire will be securely stored at Massey University; no names or identifying information is recorded. Participation in the questionnaire is anonymous, and the information will only be viewed by me, the researcher, my supervisor, and the computer programmer in the School of Psychology, Massey University.

If you have any further questions please contact the researcher or chief supervisor. If you would like to receive a summary of the research findings following completion, please send an email to the researcher at amberfletcher6@gmail.com, and one will be provided to you as soon as it is available.

Please note this project has been evaluated by peer review and judged to be low risk. Consequently, it has not been reviewed by one of the University's Human Ethics Committees. The researchers named above are responsible for the ethical conduct of this research.

If you have any concerns about the conduct of this research that you wish to raise with someone other than the researchers, please contact Professor John O'Neill, Director,

Research Ethics, telephone +64 6 350 5249, email: humanethics@massey.ac.nz

[Please CLICK HERE if you would like to continue and participate in this research](#)

Instructions

Dear Clinical Psychologist/Intern/Student

The following exercise has been designed to investigate the way in which therapists expect clients to make progress during the course of therapy, the degree of change that can be expected, and the degree to which lasting change might be seen at different lengths of follow-up.

Two typical clients are described in some detail and a fairly standard protocol of CBT intervention is suggested. Your task is to read the scenario carefully and then enter on the attached graphs specific likely data points for three possible indices that will be evident after each session.

For each client it is suggested that the therapist intends to monitor three outcome measures. You may not be entirely familiar with these measures, so some basic information is provided regarding clinical norms. Because there are three measures per client, we have provided three separate scales on the Y axis; please use these scales for each of the three measures.

It is common for clinicians who are monitoring client change to keep track of progress (or otherwise) by entering scores graphically so that general trends can be seen visually right away. When you record each data point on the electronic task you will be able to see the trends you are predicting and you can change your estimate to reflect the trends you expect any time before you submit your graphs electronically. Once you indicate you have completed the task we can record your estimates digitally.

A few additional demographic and practical questions are presented and these details will be helpful in the analysis of the data you have provided. Please take the time to fill out the questionnaire section as thoroughly as possible.

This is not an evaluation of your clinical skills! There are no right or wrong answers; we are interested in your own predictions or projections. But do assume that *you* are the therapist!

Note: If the graph fails to display correctly (bits cut off or missing) please try adjusting the Zoom level of your web browser using **Ctrl +** and **Ctrl -** (or **View ⇒ Zoom** menu in Internet Explorer and Firefox).

THANKS VERY MUCH FOR YOUR PARTICIPATION IN THIS STUDY!

IT IS APPRECIATED.

AMBER FLETCHER AND IAN EVANS

Click Continue when you are ready to proceed.

Client 1: Mr T

Client 1: Mr T

Mr T is a 35-year-old married man with two children who is employed as a manager for a supermarket chain. His GP has diagnosed clinical depression, basing this on his low mood, loss of appetite, poor sleep, inability to concentrate, in addition to finding it increasingly hard to find things enjoyable, including activities he previously liked doing. These issues have been troubling the client for approximately five months, and Mr T's GP suggested psychotherapy rather than medication. This is Mr T's first time attending psychotherapy.

Mr T is finding it increasingly difficult to go to work and he is not getting routine tasks done on time. During your initial interview (Session 1) you felt that rapport had been easily established. When treatment goals were being discussed, Mr T stated that he hoped to experience less negative moods on a day-to-day basis, to have a healthier, happier lifestyle (including the reduction of depressive symptoms), and to be able to get out and do more enjoyable things with his family or by himself, the way he used to do in the past.

As his therapist you decide, therefore, to monitor progress by routine psychological measurement. You decide to measure: (a) **mood** by means of the **Negative Affect scale of the PANAS** ; (PANAS-NA; based on the past few days, not right at this moment); (b) clinical **symptoms of depression**, using the **Beck Depression Inventory (BDI-II)**; and (c) a self-report of the **number of positive activities** he engaged in per week, based on his recall. Positive activities are defined as activities clients themselves identify as pleasurable. Depressed clients, like Mr T typically lose pleasure in activities they previously enjoyed. Most people should be able to think of 7-10 enjoyable activities that they have done in a given week, and some people will report many more than that. Together, in collaboration with the therapist, Mr T came up with a list of ten activities that provide some level of enjoyment for him. These activities included spending time with family, playing golf, and cooking.

During the first session you explain to Mr T that it will be a helpful guide to treatment if you obtain this psychometric information each week. Before the end of the session you gather data from the three measures. His Negative Affect (PANAS-NA) score was 26, which is quite a high score as a more typical score for an average adult is about 14. His BDI-II score was 35, indicating a severe level of depression—scores for people with little or no depression are usually less than 13, which is considered a clinically significant cut-off score. In terms of the number of pleasurable activities (No of Activities) he could recall in the past week (excluding going to work and watching TV), his answer was none (zero).

Your third-party payer (insurance coverage) has determined that a maximum of 12 sessions can be covered, but that you could have two follow-up or booster sessions 3 months and 6 months after the end of treatment.

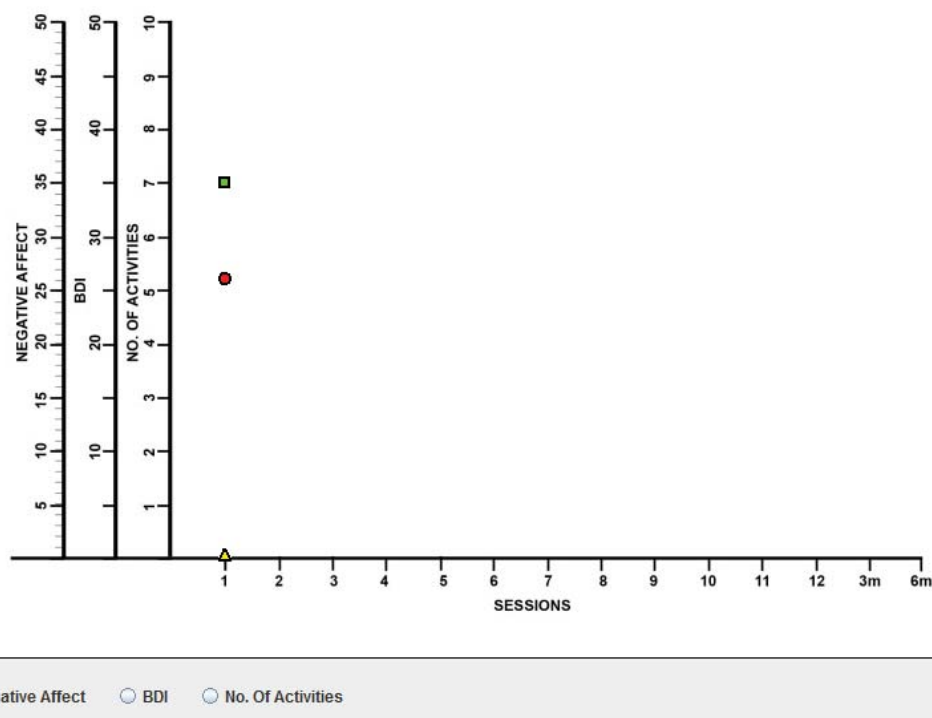
During your first session you felt that you connected well with Mr T and were positive in affirming that depression was not uncommon and typically responded well to CBT treatment.

Your treatment followed a standard evidence-based CBT protocol for depression, and included (a) some psycho-education, in which you explained some of what we know about clinical depression and how mood is related to patterns of thought; (b) cognitive restructuring of irrational thoughts, particularly focusing on the "cognitive triad" of thinking himself as worthless, the world as unfair, and the future as hopeless; (c) using role playing and Socratic questioning; (d) homework assignments in which the client was asked to notice times he had automatic negative thoughts and to write down alternatives; (e) teaching coping skills, such as problem solving; and (f) encouraging him to engage in more rewarding activities by setting goals and establishing activity schedules (often called "behavioural activation").

At the end of Session 4 Mr T commented that he had made a commitment to "turning my life around." At the end of Session 6 you felt that there was something of a breakthrough in treatment as Mr T revealed significant insight into the possible causes and nature of his problem.

For the purpose of this task, please assume that the first 12 sessions are spaced one week apart.

On the graph below, please enter your best prediction of what Mr T's scores on the three clinical measures would be for **each** of the 12 sessions (we have filled in the actual scores obtained during Session 1) and what you think you might have found at the 3-month and 6-month follow up visits.



Client 2: Ms S

Client 2: Ms S

Ms S is a 31-year-old woman with two young children; she is unmarried but in a stable relationship with her partner who is a builder. Ms S is currently at home looking after the children; before her first child was born she worked as a receptionist at a dry-cleaning business. Ms S was referred by her GP for chronic anxiety. This is her first time seeing a psychotherapist.

From the description her doctor provided it seems clear that she has generalised anxiety disorder (GAD) with some panic or agoraphobic-like symptoms (heart palpitations, feeling dizzy, thinking she was going to pass out, sweaty hands, and hyperventilation). Symptoms had developed in varying forms over the past six months. She is experiencing discomfort when at social events such as parties, when out shopping at the mall or a crowded supermarket, and is avoiding many such activities, which is disruptive for the family. When taking the children somewhere in the car she was especially nervous.

During your first session you asked her what her goals for therapy were. She stated that she wanted to be free of the symptoms of anxiety, to stop worrying and feeling nervous all the time, and to be able to engage in ordinary everyday social activities like she used to be able to do without avoiding them or always trying to escape.

Given these goals you decided to monitor on a weekly basis her **anxiety symptoms**, using the 21-item **Beck Anxiety Inventory BAI** (but asking about the past week, not the past month), her **worrying** during the week, using the 16-item **Penn State Worry Questionnaire PSWQ**, and her **avoidant behaviour** using a self-recorded daily diary (% of Activities). The diary consisted of a list you and she constructed of all the everyday activities and events she felt she *should* be able to do or go to, ranging from easy (getting bread at the convenience store), to harder activities (going to parent day at her daughter's preschool). There were two columns to be ticked off each day: one column as to whether she wanted to or needed to engage in the activity, and the second column whether she did the activity or avoided it. The weekly score for this daily event diary was the *percentage* of activities (% of Activities) on the list that she wanted to do that week that she actually did do. Some weeks the number of activities she needed to engage in could be quite small, say only 5; if she succeeded by doing (not avoiding or escaping) one of them, her score was 20% for that week. Some weeks she might list 12 things she wanted to be able to do and that she did 6 of them—that week her score would be 50%.

During your first session you explained these weekly measures and asked her to complete the three measures for the past week before she came to see you. Her score on the BAI was 48 out of a possible high of 63. Low or typical anxiety levels usually score between 0 and 7. Ms S's initial score on the PSWQ was 73 out of a maximum of 80. Someone with no reported worries at all obtains a score of 16; and someone who indicates mild or rare levels of worrying might score 32 on this scale. Ms S thought of about 11 activities or events she probably should have done or wanted to do the previous week and she did none of them, so her percentage score on the Daily Events Diary was 0.

During your first session you felt that you connected well with Ms S, although she was quite nervous. You were positive in affirming that anxiety was not uncommon and typically responded well to a CBT treatment protocol.

You introduced a standard treatment protocol. This consisted in a number of components: (a) psycho-education on the nature of fear and anxiety and what we mean by anxiety sensitivity; (b) you explored the relationship between fear and avoidance behaviour; (c) you taught her deep muscle relaxation and how to take calming breaths—practicing these when home alone or as a strategy when feeling nervous; (d) you established a hierarchy of activities from the easiest to the most difficult, encouraging her to begin with the easier ones and rewarding herself in small ways for each success; (e) you explored her catastrophic irrational thoughts by compiling a list of the possible negative outcomes of events and activities and you encouraged her replace these with more positive thoughts; (f) and you gave her some "mindfulness" exercises in which when she was feeling anxious, instead of fighting the feeling, she was to stay with it for a few moments.

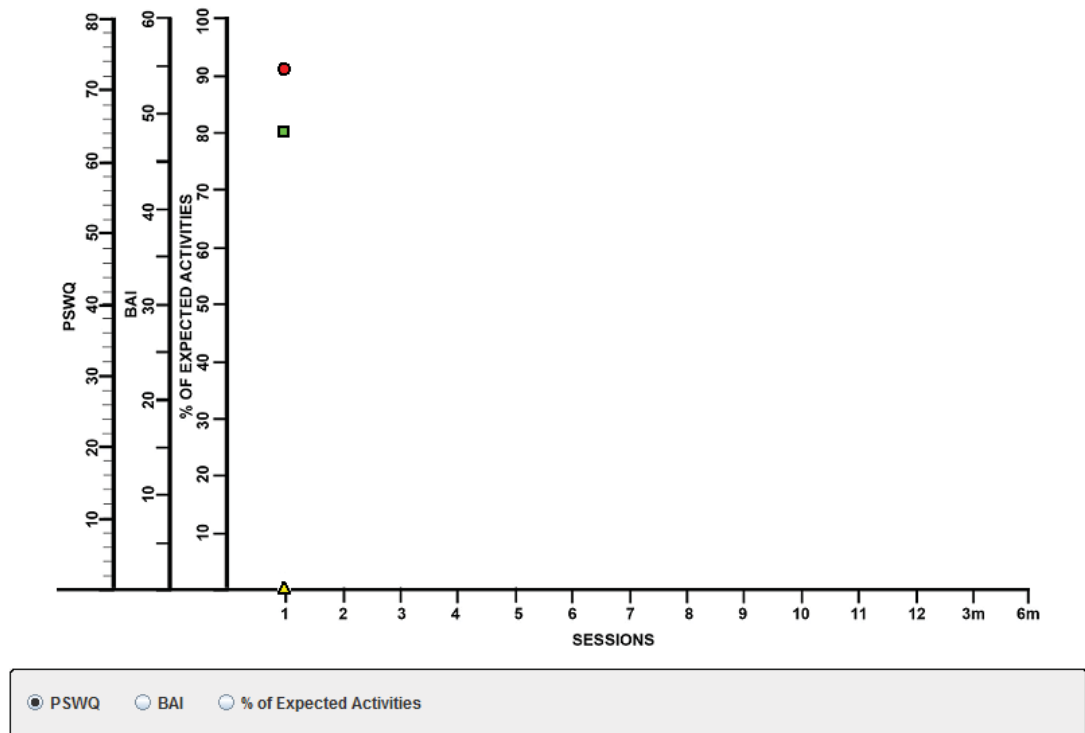
Her insurance coverage allowed you to see her for only 12 sessions, but that you could have a follow-up session at 3-months and 6-months post treatment.

At the end of Session 4 Ms. S reported an increase in self-efficacy, saying that for the first time she thought she would be able to eventually carry out all the things she had been

avoiding. During Session 6 Ms S described how she had been talking to her partner and during the discussion she had a realisation for the first time that much of her anxiety started during high school. Her parents, who were now deceased, had always been quite punitive and very critical and as a teenager she started to feel she could never live up to their expectations in terms of school work, looks, personality, and so on.

For the purpose of this task, please assume that the first 12 sessions are spaced one week apart.

On the graph below, please enter your best prediction of what Ms S’s scores on the three clinical monitoring tasks would be for **each** of the 12 sessions (we have filled in the actual scores obtained during Session 1) and what you think you might have found at the 3-month and 6-month follow up visits.



Questionnaire

Now that you have **finished** the task, would you please answer the questions below:

1	Are you a student	<input type="radio"/> Yes <input type="radio"/> No
2a	If yes, what year of the clinical programme are you currently in?	<input type="text"/> years
2b	If no, how many years of clinical experience do you have?	<input type="text"/> years
3	What is your profession (job title)?	<input type="text"/>
4	What country are you from?	<input type="text"/>
5	How many of the measures described are you familiar with?	<input type="radio"/> All <input type="radio"/> Some <input type="radio"/> None
6	I have used CBT with anxiety disorders.	<input type="radio"/> Yes <input type="radio"/> No
7	I have used CBT with depression.	<input type="radio"/> Yes <input type="radio"/> No

Are you aware of any models/theories in predicting client change? Yes/No

If yes, which models/theories are you aware of?

Thinking of the process you went through in predicting the shape of change for each client, were there any ideas/concepts/theories that you can recall that impacted on your expectations? i.e. “this is a complex case and I’d not expect much improvement” or “I’ve used the BAI and scores seem to vary from week to week.” We are interested in your reasoning about monitoring client progress.

Do you have any other comments you would like to make about this task?