Client and Therapist Variability with Psychotherapy Homework: A Preliminary Psychometric Evaluation of Two Scales

A thesis presented in partial fulfilment of the requirement for the degree

of

Master of Arts

in

Psychology

at Massey University,

Albany

Margo-ann D. Munro

2006

Abstract

The therapeutic use of homework assignments is an integral feature of Cognitive Behaviour Therapy (CBT) and is believed to be important in producing and maintaining client treatment gains. While there is empirical and theoretical support suggesting that homework completion may be associated with improved outcome, few studies have directly investigated specific therapist behaviours used in the integration of homework into practice. Similarly, research on client homework completion has, by far, focused solely on the quantity of homework completion rather than on any other factors found to be associated with increased client engagement with homework assignments. The recently developed Homework Adherence and Competence Scale (HAACS) was designed to specifically assess therapist competence and adherence to recommended homework behaviours used to administer homework. As a new measure of client homework completion, the Homework Rating Scale-II (HRS-II), extends the assessment of client homework completion beyond quantity only and provides a measure of the cognitive and behavioural determinants of client homework completion. The present study presents a psychometric evaluation of both the HAACS and HRS-II when rated by independent observers. Further, the present study sought to examine therapist differences in adherence and competence in administrating homework as well as investigating the temporal pattern of client homework completion.

ACKNOWLEDGEMENTS

I would like to acknowledge Massey University for their funding contributions that made this project possible. My supervisor's research grant that funded the raters in this study, and my Massey Scholarship, which paid for the tuition in my paper year.

I would like to commend Dr. Nik Kazantzis for his extreme patience and understanding throughout the course of my thesis. Through all the trials and tribulations, his support, enthusiasm and commitment to the project has made it worthwhile. I would also like to extend thanks to Dr. Keith Dobson from the University of Alberta for his generosity in allowing me to have access to the DVDs and CTS ratings that formed the basis for my project. His feedback throughout the project has proved invaluable.

A huge thanks to Angela and Tracey, my raters who spent innumerable hours watching and rating the DVDs for the study. Their continued enthusiasm for the project made it all worthwhile and provided me with a greater understanding of the back roads in western and central Auckland. Also, I would like to thank the current and past members of the research lab group for their input and ongoing support for my project. I would especially like to thank Paul. Judith, Gregg and Aniel who have supplied priceless guidance regarding the process of completing a thesis.

I would also like to acknowledge Dr. Barry MacDonald, who supplied great statistical support when I needed it most. I could not leave out the library staff, especially Joy Oehlers, whose help in securing and extending book loans was greatly appreciated. As well, a big thank-you to the folks in IT who helped me solve my Mac "incompatibility with library databases" problem.

I would like to thank everyone who played a part in the ongoing saga, which was my thesis (I'm seriously thinking of giving it all up to go write Country and Western songs, as I think I have enough material to succeed!).

I thank the members of the Solo Mums Club for their help in supplying childcare and anti-stress techniques when needed it. I thank Tracey and Colin, Nancy and Paul and all the others who provided a shoulder and an ear when I needed it. As for Liz, well, thanks for all the coffee and for running the batteries dead on my phone as we discussed our theses. I look forward to finally celebrating their completion with you.

Finally, and most importantly, I would like to acknowledge my family for their ongoing support and patience throughout this project. To my parents, who have phoned every week from Nova Scotia, Canada to see how it was going: thank you for being there, if not in person. To my son Ian: thank you for being such a great kid, I love ya heaps... ZOZOZO!

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

Introduction

Overview

The following chapter provides an overview of the support for the use of homework in Cognitive Behavioural Therapy (CBT), along with the primary aims and rationale for the present study. Within this overview, the concepts of homework compliance and therapist adherence and competence are defined. The placement of this research as part of the ongoing Cognitive Behavior Therapy Research Project is summarised, as will the chapter outline for this research.

Introduction

Within the realm of education, homework is not an alien concept. Although school children may not necessarily appreciate time away from out of school activities, the premise of homework is to consolidate and extend what is taught in the classroom. Educators often employ a technique known as 'scaffolding' to engender self-learning in students; by applying most instruction in the early stages of learning a new skill, the basic building blocks are laid upon which the student is able to add-on, extend and generalise into other areas of the curriculum (Berk, 2001).

Similarly, within psychotherapy, homework tasks can be viewed as a way to generalise in-session practice into the real world and has long been a core feature of Cognitive Behaviour Therapy (CBT; A. T. Beck, Rush, Shaw & Emery, 1979). The objective of the therapeutic use of homework is to reinforce in-session learning and to apply this into natural settings (Tompkins, 2003).

While there are similarities regarding the rationale for homework within both education and psychotherapy, there are some differences. Homework assigned by a teacher is generally meant to aid academic achievement, through review, advance preparation or elaboration of previously learned material and once completed, a student may relax until the next homework is assigned (Hong & Milgram, 2000; Wood, 1998). This highlights the main difference between academic homework and therapeutic homework; therapeutic homework extends the work in the therapy session throughout the the hours of the week, in what has been described as providing multiple sessions for the price of one (Beck, 1995) Further, Detweiler and Whisman (1999) posit that the therapy session is a small portion of the time that a client lives with their problems, and homework assignments help to address that imbalance by providing a means to enhance newly learned skills. That is, the main rationale for the use of homework in CBT is the provision of an avenue for patients to practice and generalise newly learned skills within their natural environment (Persons, Davidson, & Tompkins, 2001), Further, the structure that homework provides for the practice of skills, may be useful in relapse prevention, as homework aids clients in identifying problems and use skills learned in therapy to solve these problems (Detweiler & Whisman, 1999; Persons, 1989; Simons, Murphy, Levine, & Wetzel, 1986).

Empirical Support for Homework

While homework is a core ingredient of CBT (Beck, et al., 1979), a recent review of research literature published between 2000 and 2005 found very little empirical evidence for a causal relationship between homework completion and improved outcome (Kazantzis, Deane, Ronan, & Lampoproulos (2005). Further, although there is a great deal of published research including homework assignments

as part of the prescribed protocol (for reviews, see Kazantzis, 2000; Mahrer, Nordin, & Miller, 1995), research findings have not always been conclusive (Kazantzis & Lampropoulos, 2002). Neiymeyer and Feixas (1990) conducted a randomised clinical trial using Group CBT in order to examine the influence of homework on outcome. The sample consisted of outpatients meeting the criteria for major depressive disorder (MDD). Patients were randomly assigned to Group CBT conditions that either used or did not use homework assignments as part of their protocol. While patients in both conditions improved significantly on outcome measures of depression, the Beck Depression Index (BDI; Beck, Ward, Mendelson, Mock, & Erbaugh, 1961) and Hamilton Depression Rating Scale (HAM-D; Hamilton, 1960), Time X Treatment interactions were not significant, indicating that the systematic use of homework was not associated with improved outcome. However, post hoc tests, accounting for pre-treatment severity found that patients using homework improved significantly more on the HAM-D than did those not using homework. This finding was not replicated for scores on the BDI. Based on these findings, the authors suggested that homework facilitated recovery for more depressed patients than those patients with milder depression.

There has been some discussion as to methodological limitations in studies that may provide some explanation as to why the results pertaining to the efficacy of homework has been mixed. Kazantzis, Dattilio and Merrick (in press) reviewed psychotherapy outcome research published between 1993 and 2003 and found that while 64 percent of the reviewed studies reported the use of homework assignment as part of the protocol, only 6% of those studies included an assessment of client homework compliance. As noted by the authors, this has definite implications as to

the internal validity of these studies, as homework non-completion may confound findings of non-significance related to homework effects.

Further, it has been suggested that early studies assessing homework effects were not design with enough power to detect an effect (Kazantzis, 2000). In a power analysis of 27 studies, Kazantzis found that, on average, studies were designed in such a way that there was only a 32% chance of detecting a significant effect. That is, studies are not designed sensitive enough to detect an effect (Kazantzis & Lampropoulos, 2002). It should be noted, however, that the issue of low power to detect significant effects is not limited to the study of homework. In a recent review of the detection of therapist effects, Crits-Christoph and Gallop (2006) suggest that, at best, large multi-site studies will only be able to detect small effects for identifying therapist differences and suggest that the best option for increased power is large scale naturalistic studies (i.e., Okiishi, Lambert, Nielsen, & Ogles, 2003).

While reviews of published research literature reveals ambiguity as to the effect of homework compliance on outcome, research in CBT specifically focussed on homework has sought to examine whether quality or quantity of homework completion has a significant effect on outcome. As a core ingredient in the delivery of CBT (A. T. Beck, et al., 1979, J. S. Beck, 1995), the clinical utility of homework assignments is of particular interest to CBT practitioners. The mainstay of research on the client's role in treatment outcome (i.e., compliance with homework recommendations), has been undertaken in the study of the treatment of depression (see for example, Burns & Nolen-Hoeksma, 1991; Fennel & Teasdale, 1987; Neimeyer & Feixas, 1990; Startup & Edmonds, 1994). In the main, there is evidence supporting the proposition that depressed clients who complete a large quantity of homework, and do it well, are likely to have improved outcomes when compared to clients who

do not. However, when homework completion is assessed in anxious, rather than depressed, clients, findings suggest that quantity and quality of homework completion is not associated with better outcome as assessed by outcome measures of anxiety (e.g., Edelman & Chambless, 1995; Lampoproulos & Rector, 2004; Woody & Adessky, 2002). There are a few notable exceptions to this finding; Park, Mataix-Cols, Marks, Ngamthipwatthana, Marks, Araya, et al. (2001) and Schmidt and Woolaway-Bickel (2000) did find significant associations between homework completion and reduced anxiety. Further, Edelman and Chambless (1995) found that although homework had no immediate effect (that is, by the end of treatment), clients with social phobia who were rated as homework compliant reported the largest decrease in both anxiety and avoidant behaviour at 6-month follow-up.

Research in CBT has suggested that early compliance with homework is associated with improved outcome (e.g., Fennell & Teasdale, 1987). Activity scheduling, a central component in Beck's (A. T. Beck et al., 1979) treatment for depression, is often one of the early homework interventions. According to Beck, the symptoms of depression consist of mood, behaviours and cognitions. These three factors interact in a reciprocal causal manner, and activity scheduling can help alleviate negative mood by providing direct disconfirming evidence of their negative automatic thoughts (A. T. Beck et al., 1979; Persons, Davidson, & Tompkins, 2001). Activity scheduling is useful in promoting a sense of mastery and pleasure, while providing a behavioural experiment to test out negative assumptions (Persons, Davidson, & Tompkins, 2001). Further, Wright, Thase, Beck, & Ludgate, 1993) suggest that simply becoming active and engaging in homework provides reinforcement, which in turn, increases a client's motivation to complete homework.

This premise is similar to Lewinson's rationale for the use of pleasant activity scheduling in behaviour therapy (Hoberman & Lewinsohn, 1985).

While findings regarding the empirical support for homework may differ across presenting problems, the underlying rationale for the use of homework is that it provides an opportunity for client learning which is consistent with CBT's learning approach to therapy (Persons, 1989). Kazantzis, Deane, & Ronan (2005) propose that the learning derived from homework assignments is key in understanding the quality of homework completion.

Definitions

Homework can take many forms, from activity scheduling, bibliotherapy, information gathering exercises, to monitoring thoughts, feelings and behaviours. Just as there is any number of out of session activities, which would constitute being considered homework, there are just as many terms used to describe homework (Kazantzis & Ronan, 2006). In the context of the present research, *homework* is defined as any task that contributes towards the gathering of information by the client, is directly associated with the goals of therapy, are completed outside of the normal session time and allow the client to generalise in-session learning into their natural environment (see for example Alford & A. T. Beck, 1997; A. T. Beck, Rush, Shaw & Emery, 1979; Lambert, Harman, & Slade, in press).

Homework compliance is another term requiring a definition within the context of this research. As discussed earlier in this chapter, the assessment of client compliance has generally focussed on the quantity or quality of the homework completed by a client. In a brief review of 32 studies that described homework assessment, Kazantzis, Ronan and Deane (2004) identified that the assessment of

homework compliance is predominantly based on the degree to which a client complied with the assignment of homework. The authors also noted this assessment is usually made without consideration of how well the homework was completed, or other factors that may mediate or moderate homework completion. The authors later proposed that compliance encompasses a much broader set of factors referred to as homework completion (Kazantzis, Deane, & Ronan, 2005). Thus, *homework compliance*, as assessed in the present study is defined as homework completion. In brief, homework completion is the benefit, or learning, gained from homework assignments. Involved in this is the interaction of client beliefs about the task that may help or hinder a client to overcome potential obstacle that may arise in relation to doing the task. Further, completion relates to the amount of homework completed, the extent of learning gained from doing the homework, and the resulting consequences and conclusions derived from the experience (Kazantsis, Deane, & Ronan, 2005). The model for assessing homework completion is described in Chapter Three.

When research is undertaken, it is important for both the internal and external validity of the study that therapist behaviours be monitored for both adherence to the protocol and the skill with which these behaviours are delivered (Perepletchikova & Kazdin, 2005). The overarching issue is treatment integrity (or fidelity), which refers to the extent to which a treatment was provided as intended. The need for the assessment of treatment integrity is not limited to research applications, but has direct implications for training and clinical supervision (e.g. Sudak, J. S. Beck, & Wright, 2003; Pretorius, 2006). In the context of this study, *therapist adherence* is defined as the degree to which a therapist follows recommendations for the integration of homework into practice. *Therapist competence* is defined as the degree of skill shown when performing these behaviours.

The preceding paragraphs raise a paradox; homework is seen as a crucial element in CBT, yet its association with treatment outcome is tenuous. First, there are measurement issues, which may be associated with the unclear results. As noted, most measures of homework compliance focus on the quantity of homework completion rather than assessing the quality of completion and the client's experience (for a review, see Kazantzis , Deane & Ronan, 2004). Second, studies have provided minimal evidence of therapist adherence and competence in administering homework. The measures for therapist homework behaviours tend to be assessed on the basis of one or two items from overall therapist competence measures. Perpletchikova and Kazdin (2005) suggest that specific measures need to be used to assess therapist competence and adherence, and that these two facets of treatment integrity should be assessed separately in order to reduce the possibility of confounding the two (Barber & Crits-Cristoph, 1996). Further, Doss suggests that the identification of specific therapist change processes (i.e. relevant behaviours) may be an important first step in the understanding of client change mechanisms.

Cognitive Behavior Therapy Homework Project

The overall aim of the Cognitive Behavior Therapy Homework Project is to develop an understanding of specific mehanisms by which homework produces its effect within CBT. Specifically, the project has five broad ojectives:

1. To carry out statistical and conventional reviews of empirical literature in order to consolidate and clarify current understanding of homework.
2. To survey practitioners' use of homework assignments in clinical practice in order to determine the necessity, utility and areas for future research.

3. To design a theoretical model and treatment manual for the integration of homework assignments into practice.
4. To design conceptually driven measures of therapist competence in homework integration and client homework completion. Part of this aim is the psychometric evaluation of the new measures.
5. To conduct prospective process and treatment outcome research in order to evaluate the utility of the theoretical model and treatment manual developed by the Cognitive Behavior Homework Project.

Within this context, the present study contributes to the fourth objective of the research project, namely provide an evaluation of the psychometric properties of two measures developed by the team research project. This psychometric evaluation is based on the use of these measures by independent raters to assess therapist adherence and competence in the integration of homework as well as the assessment of client homework completion. The present study used archived data from an existing randomised clinical trial assessing the effectiveness of Behavioural Activation, Cognitive Therapy, Antidepressant Medication and pill placebo (Dimidjian, Dobson, Kohlenberg, Gallop, Markley, Atkins, et al., in press). Cognisant of this fact, the discussion includes specific recommendations regarding the generaliseability of the findings of this study and includes recommendations for future research methodology, which will help overcome methodological issues that arose in the present study.

Chapter Outline

Chapter 2 describes cognitive and behavioural theories of learning as they relate to the learning focus of homework. Following on from learning, Chapter 3 provides theories of motivation associated with increasing client compliance with homework assignments. In addition, Chapter 3 presents recommendations for the integration of homework into practice, which have been associated with increasing client compliance. Following on from these recommendations is the assessment of homework compliance and the development of a theoretically driven measure of client homework compliance, the Homework Rating Scale (HRS-II, Kazantzis, Ronan & Deane, 2005).

Chapter 4 describes the need for the assessment of therapist competence and adherence in relation to treatment integrity, clinical training and ongoing clinical supervision. The assessment of overall therapist competence and adherence in CBT is then described, followed by the extension of therapist assessment solely related to homework integration.

As the present study provides a psychometric evaluation of two new measures, Chapter 5 outlines recommendation regarding the gathering and reporting of psychometric evidence. Included in this chapter is a review of three recent psychometric evaluations of new or modified measures.

Chapter 6 provides a brief summary of the implications from chapters two through five, as they relate to the context of the present study. The research aims are stated in full, along with specific hypotheses.

Chapter 7 presents the method for Study One, the psychometric evaluation of the HAACS. The rater training procedures are described as are the statistical analyses used in this study. Chapter 8 presents the results of Study One. Results are separated

into the specific areas of evaluation. Estimates of the reliability of the ratings are presented first, with adherence and competence ratings assessed separately. Following on from this, the results of the assessment of the internal validity of the HAACS are presented. Next, results of the correlational analysis between the HAACS and other measures of therapist competence and/or adherence are described. The chapter ends with the results of specific analyses directed at evaluating therapist differences in adherence and competence in the administration of homework. Chapter 9 presents the discussion of the main findings for Study One.

Chapter 10 presents the method section for the second study, the psychometric evaluation of the HRS-II. Included is the description of the measures used in Study Two as well as the statistical procedures relevant to that study. The results for Study Two are presented in Chapter 11. Again, results are presented according to the specific analyses involved. The reliability estimates for the ratings are presented first, with the results of the correlational analysis between the two measures used in Study Two. Following from this are the results from the Bland-Altman analysis of the levels of agreement between the two measures. The results of the assessment of the internal validity of the HRS-II using Principal Components Analysis are presented next, including a reliability analysis of the factor structure. The resulting factors are then compares to the factors derived using therapist ratings and client ratings from another study using the HRS-II. Finally, the results from the analysis of differences in homework completion across therapy are described. Chapter 12 provides a discussion of the main findings from Study Two.

Chapter 13 provides an overall general discussion, including a summary of the limitations of the present study, the significance of the study, and contribution in terms of future clinical and empirical applications.

CHAPTER 2

Homework in Psychotherapy

Overview

Homework in school is not a new phenomenon, neither is the use of out of session tasks within psychotherapy a new phenomenon. Just as school homework is not specific to one subject and not others, therapeutic homework is not specific to one theoretical paradigm (Kazantzis & Deane, 1999). Early documentation of the benefits of between-session tasks stems from the psychoanalytical paradigm. Homework, albeit not referred to as such, has been identified as being advantageous in overcoming neurosis (Herzberg, 1941), with client engagement with graduated exposure activities between sessions resulting in improved rate of progress.

Across therapeutic paradigms, homework is viewed as a means by which the effectiveness of therapy is improved (e.g., Greenberg, Watson & Goldman, 1988; Masters & Johnson, 1970). Homework is increasingly being utilised in psychotherapy, regardless of the theoretical orientation of the therapist, with most psychologists reporting the use of homework tasks within their clinical practice (Kazantzis & Deane, 1999; Kazantzis, Lampropoulos, & Deane, 2005). The integration of homework tasks has been identified in numerous approaches to psychotherapy including behavioural (Martin & Pear, 2004; Shelton & Ackerman, 1974), cognitive (Beck, 1964), emotion-focused therapy (Elliot, Watson, Goldman, & Greenberg, 2003; Greenberg, Watson, & Goldman, 1988) and systemic family therapy (Datillio, 2002).

The underlying rationale for the specific use of homework tasks in therapy is underscored by each paradigm's particular stance regarding theories of learning. This chapter will look, in particular, at the main behavioural and cognitive theories of

learning underpinning the use of homework within the cognitive behavioural framework (A. T. Beck, Rush, Shaw, & Emery, 1979). Firstly, the principles of classical conditioning, operant conditioning, and generality supply the behavioural perspective regarding learning will be described. Secondly, the cognitive viewpoint, including social learning/social cognition theory, will be addressed. Together, these theories help describe the underlying learning principles associated with homework.

Behavioural Theories

The rationale for the use of homework in cognitive behavioural therapy has evolved from various theories of learning. Learning is a process through which durable change in behaviour, knowledge or cognition is derived through experience (Martin & Pear, 2003). Not only does learning include the acquisition of skills and knowledge, but is behind the shaping of enduring personal habits and emotional responses. According to the behavioural viewpoint, the principles of classical and operant conditioning results in learning at a fundamental level based upon associations made through interaction with one's environment.

Classical Conditioning

Classical conditioning, or Pavlovian conditioning, refers to learning in which a response originally evoked by one stimulus is acquired by another stimulus (Martin & Pear, 2003). At a very basic level, there is an innate, unconditioned association between certain stimuli (*unconditioned stimuli*) and responses (*unconditioned responses)*. In this instance, a response is naturally occurring, and not learned or conditioned. Typically, these responses are involuntary reflexes and visceral responses (Martin &

Pear, 2003). On the other hand, a conditioned association occurs when an unrelated and previously *neutral stimulus* acquires the capacity to evoke a *conditioned response*. In this instance, the response is a learned, or conditioned, response. The pairing of a previously *neutral stimulus* and *unconditioned stimulus* may result in the neutral stimulus acquiring the ability to evoke an unconditioned response. For example, while driving a car, an individual almost has a car accident, which elicits an unconditioned fear response. If, however, at the same time there was a Celine Dion song playing on the radio, this may become paired with the fear elicited by the near accident. Once this association is made, the mere sound of Celine Dion song may be sufficient to evoke a fear response in the future where this had not previously occurred (Rachman, 1990). In this instance, the neutral stimulus (Celine Dion song) would now be a conditioned stimulus and the fear response conditioned to that stimulus.

Classical conditioning is often indicated in the formation of emotional responses such as fear (Martin & Pear, 2003). Irrational fears associated with phobias can often be traced back to experiences involving classical conditioning (Ayres, 1998). Fortunately, a conditioned response tendency can be weakened and/or disappear through extinction (Martin & Pear, 2003). According to classical conditioning, when a conditioned stimulus is consistently presented alone without the unconditioned stimulus, the capacity for the conditioned stimulus to evoke the conditioned response is decreased. For example, an individual tenses up whenever they hear a Celine Dion song, which had been paired with fear in the past. This individual may work with their therapist to extinguish this response through listening to Celine Dion music. The presentation of Celine Dion songs without the original unconditioned stimulus would gradually decrease the conditioned response to the point of extinction.

Extinction underlies a number of behavioural focused homework tasks, such as flooding and systematic desensitisation (Kazantzis & L'Abate, 2005). The utility of homework tasks based on the principle of extinction lies with the ability of the client to engage with the extinction process in a natural, real world setting. This helps to ensure that the extinction of the stimulus-response generalises into other situations and is not limited within the therapy environment (Kazantzis & L'Abate, 2005)

Operant Conditioning

Whereas classical conditioning generally produces reflexive responses depending on a preceding stimulus, operant conditioning is a form of learning based upon behaviour, or responses, occurring as a result of their consequences (Martin & Pear, 2004). Operant conditioning, or instrumental learning, is based on behaviours being contingent on an expected consequence (Berk, 2001; Martin & Pear, 2003). That is, positive consequences reinforce behaviour, whereas negative consequences tend to reduce the likelihood of future occurrence of behaviour.

Behaviour therefore, can be shaped through the manipulation of reinforcement (Marin & Pear, 2003; Thorndike, 1927). In terms of operant conditioning, behaviour develops through the expectation of benefits or rewards. In panic, avoidance is often a behaviour resulting in the reduction of anxiety, which reinforces avoidance as a tenable option in the future (Barlow & Durand, 2002; Persons, 2001). However, while there may be short-term benefit in the use of avoidance as a strategy to reduce anxiety, long-term consequences may result, with

these avoidance patterns becoming ingrained and extended into other areas of functioning (Persons, 2001).

Just as behaviour develops through the expectation of reinforcement, behaviour can also be changed using the same principles (Martin & Pear, 2003). In order to increase the likelihood of client homework completion, a therapist can use various forms of reinforcement in the design of homework tasks (Kazantzis & L'Abate, 2005). The use of reinforcement increases the chance that behaviour will be initially enacted and then repeated in the future (Padesky & Greenberger, 1995). However, determining which reinforcers will work can prove problematic for a therapist, as an individual's reinforcers are idiosyncratic and have developed within the context of their own environment (Martin & Pear, 2003).

While therapist reinforcement (e.g., praise and encouragement) may initially help with client engagement, reinforcement in natural settings may serve to maintain and to generalise behaviour (Kohler & Greenwood, 1986). Further, the chance for a client to experience reinforcing responses increases when behaviour is enacted within a real-world setting (DeLeon & Iwata, 1996). Similarly, the probability of behaviours being reinforce quickly is increased when behaviour is enacted within the home and work environments (Kohler & Greenwood, 1986).

Reinforcement for behaviour is not always extrinsic, or external. Engagement with behaviour can result in intrinsic, or internal, reinforcement. A. T. Beck, et al. (1979) posit that behavioural consequences may be internal, such a sense of mastery or pleasure when accomplishing a homework task. Thoughts and feelings, according to cognitive theory, provide reinforcement for both external and internal behaviour (A. T. Beck, et al., 1979; Cautela & Kearney, 1993; Young, Klosko, & Weishar, 2003).

The use of specific homework tasks targeting internal behaviours, such as thought records, may provide an avenue for reinforcement that may not be obviously or externally reinforcing. Determining what is reinforcing about a homework task, be it internal or external, is a necessary part in improving homework compliance (Kazantzis & L'Abate, 2005; Kazantzis, MacEwen, & Datillio, 2005). Therefore, attention must be given to reinforcing thoughts and feelings that arise in the client regarding homework. Eliciting the client's thoughts concerning homework is useful, as it may uncover potential barriers (A. T. Beck et al., 1979; Kazantzis, MacEwen, & Datillio, 2005) to homework engagement due to negatively reinforcing thoughts regarding the task (e.g., task is too difficult). Working with a client to enhance positive reinforcing thoughts that occur following homework completion, provides ongoing reinforcement and assists the process of change for the client (Padesky & Greenberger, 1995).

Theory of Generality

Central to both classical and operant conditioning is the concept of generalisation. That is, the transfer of learning from one environment into another environment (Martin & Pear, 2004). Generalisation is an adaptive learning process, allowing for the use of previously learned behaviours in novel; without this, an individual would need to acquire completely new skills when confronted by new situations (Kazantzis & L'Abate, 2005).

Cognitive Theories

Behavioural theories of learning place a great deal of importance on environmental factors in the shaping of behaviour, cognitive theories posit that

behaviour is determined through their perceived interpretation of events. In short, beliefs, thoughts and attitudes affect behaviour and faulty thinking is the root of behavioural and emotional problems (Beck, et al., 1979). While there is no doubt that behavioural precepts play an important role in understanding the function of therapeutic homework in engendering change, cognitive processes are always involved (Cautela & Kearney, 1993).

Social Cognition

Just as there are behavioural theories of learning, there are specific learning theories within the cognitive paradigm. Social cognition theory evolved from Bandura's theory of learning known as observational learning (1977). Whereas behavioural models of conditioning posit that learning is derived from direct experience, observational learning proposes that both classical and operant conditioning can occur vicariously through the observation of another's conditioning (Bandura, 1977, 1986). Inherent in this model is the underlying cognitive mediation of behaviour.

Building upon the observational learning model, Bandura developed an interactional model of behaviour with cognitive processes central to its production. Social Cognitive Theory (Bandura,1986) views individuals as both self-reflecting and self-regulating. Inherent is the underlying notion that individuals possess self-beliefs, also referred to as self-efficacy beliefs, and these beliefs are necessary judgements of an individual's capabilities in the performance of behaviour. Thus, motivation to change, hinges upon self-efficacy (Bandura, 1986). In terms of homework, self-efficacy

regarding one's ability to accomplish a homework task will directly influence whether or not the task will be attempted for the first time or repeated in the future.

Learning, according to the social cognition model, reflects the bi-directional relationship between an individual's experience and their environment (Bandura, 1977, 1986). Thus, the individual facets of an individual's experience (i.e., cognitive, behavioural, physiological and emotional) influence the other parts of their experience. As a theory of learning, this is an all-inclusive model. In understanding the utility of homework tasks, this model posits that work in one area will result in change in others. Further, how an individual learns from homework is increased to include watching others, not simply doing, as is the case with behavioural theories. In-session practice with the therapist allows for modelling, or learning through observation, to occur (Mischel, 1973). Modelling homework tasks provides an opportunity to observe the consequences of the actions of another and form self-efficacy beliefs based upon the observation. Similarly, the use of a rationale which highlights the benefits of homework for others can influence how a client views their potential benefit from engaging in homework (A. T. Beck, et al., 1979; Persons, 1989).

Apart from modelling and providing a rationale for homework, collaboration in designing and assigning homework tasks allows the client to become an active and engaged participant with their treatment (Prochaska, DiClemente, & Norcross, 1992). When active involvement occurs within the homework process, this can increase feelings of self-efficacy (Kazantzis, & L'Abate, 2005). These self-efficacy beliefs have relevance to ongoing engagement with homework tasks; prior experience significantly influences self-efficacy beliefs (Bandura, 1986, 1989). The in-session review of the client's homework experience provides an opportunity for the therapist to gauge how

the client has synthesised their homework experience. Further, encouragement provided in feedback to the client about completed homework enhances self-efficacy beliefs as a result of social persuasion (Kazantzis & L'Abate, 2005).

Cognitive Theory

According to A. T. Beck, et al. (1979), affect and behaviour are determined by personal interpretations of events. Fundamental to these interpretations are underlying schemas (or latent constructs) regarding self, others and the world, known as the cognitive triad (Beck, 1976). According to Beck (1983), schema can be dormant, and are activated when life events match the schema. Beck's concept of psychopathology wrests upon the notion that unless a vulnerable individual experiences an event that matches and activates maladaptive schema, a person will not present with psychopathology (Persons Davidson, & Tompkins, 2001).

Originally conceived as a theory of depression, Beck (1983) posited that there were two typical types of depressed patients: the sociotropic (or dependant) and the autonomic (independent). The autonomic type views independence or success as necessary in order to be happy; loss of independence or failure to succeed may result in depressive symptoms. Alternatively, the sociotropic type views interpersonal worth as necessary to be happy; interpersonal loss or rejection may result in depressive symptoms. The identification of specific underlying schema provides evidence of Beck's two subtypes of depression (A.T. Beck, 1983; Persons, Davidson, & Tompkins, 2001). There is some evidence supporting the presence of sociotropic and autonomic subtypes of depression (e.g., Brown, Hammen, Craske, & Wickens, 1995; Imber, Pilkonis, Scotsky, Elkin, Watkins, et al., 1990).

Beck's cognitive theory of depression (1976) proposes that depression is the result of excessive dysfunctional thinking and the identification and restructuring of these dysfunctional thoughts produces changes in psychological functioning. Reality checking or hypothesis testing is often employed to determine whether or not these beliefs are indeed accurate. According to Beck (1976), distorted negative schema are learned in childhood and despite conflicting evidence later in life, these schemas remain virtually unchanged. Depression, according to A. T. Beck (1976) is comprised of mood, behaviour and cognitions (*automatic thoughts*). These cognitions are thoughts that automatically occur and are linked reciprocally with mood and behaviour. From a cognitive standpoint then, it is expected that a change in one of these aspects will produce a change in the others.

Cognitive learning theories view change as requiring self-awareness, critical analysis and synthesis (Kazantzis & L'Abate, 2005). The use of homework tasks, from a cognitive theory perspective, provides an opportunity for gathering information about thoughts, emotions and behaviours (Beck, et al., 1979; Persons, Davidson, & Tompkins, 2001). Performing homework tasks within a real-world context, allows for the client to gather data, which would otherwise not be assessable within a therapeutic session (Tompkins, 2004). That is, homework provides an opportunity for the client to test hypotheses about expected responses (A. T. Beck et al., 1979).

Summary

The therapeutic use of homework is not a new phenomenon, nor is its use limited to one specific paradigm. This chapter reviewed learning theories from the

behavioural and cognitive standpoint in order to provide a rationale for the utility of the therapeutic use of extrasession tasks. The premise behind homework is that just as behaviour is learned, it can be unlearned. Homework is, therefore, a method in which a client extends in-session learning into a real world context, thus allowing for the synthesis of the learning experience be generalised into new, or novel situations.

CHAPTER 3

Homework Completion

Overview

Not only has therapeutic homework been identified as an essential factor common to diverse therapeutic approaches to psychotherapy, the role of homework has been the subject of increased research focus, particularly the relationship between homework and treatment outcome. There have been inconsistent findings when contrasting therapies with and without homework in order to examine the specific effects of homework on outcome in CBT. Some findings report statistically significant effects on outcome associated with therapy with homework compared to therapies without (e.g., Kazdin & Masticelli, 1982).

Compliance is often used to describe whether or not an individual completes out of session tasks (Kazantzis & Ronan, 2006). This chapter will examine theories of client motivation in relation to client homework completion. Further, recommendations for the integration of homework into practice will be outlined. The issue of client and therapist factors relating to compliance with homework will be discussed, as well as the assessment of client homework completion.

Theories of Compliance

Theories regarding compliance are generally hedged within social psychology and biomedical models of health behaviour. Knowing the underlying processes that either encourages or discourages client engagement with homework provides useful tools for the clinician to employ when homework compliance is required.

Cognitive Dissonance Theory

According to cognitive dissonance theory (Festinger, 1957), discrepancies between behaviour and attitudes results in attitude shift in order to maintain consistency amongst cognitions. When two cognitions, or thoughts, are inconsistent, tension, or dissonance is the result. Festinger argued that in order to reduce this tension, adjustments in attitudes resulted. The underlying premise is that people are motivated to maintain consistency in cognitions.

Within CBT, a main focus for therapeutic intervention is cognitive restructuring (Beck, et al., 1979). This technique aims to identify and modify maladaptive cognitions and to replace these with a more adaptive way of information processing (Sternberg, 1996). However, old ways of thinking and behaving may be so ingrained, that when faced with alternative ways of thinking and behaving, a client sometimes feels cognitive-emotive dissonance or cognitive-behavioural dissonance (Beck, et al., 1979).

Difficulties in overcoming dissonance may lead a client to feel that practicing new, adaptive cognitions requires too much effort (Ellis & Dryden, 1997). It has been suggested that the primary strategy in overcoming dissonance is to strengthen the new adaptive cognitions (A. T. Beck et al., 1997; J. S. Beck, 1994). Homework tasks provide on-going, out of session practice meant to strengthen new and more adaptive cognitions and behaviours (A. T. Beck, et al., 1979).

Stages of Change

This transtheoretical model evolved from a comparative analysis of theories of psychotherapy and behaviour change (Prochaska, 1979). This model posits that motivation to change can be enhanced through therapist intervention (Prochaska,

DiClemente, & Norcross, 1992). This model proposes that there are six distinct stages, or steps, the progression through which can be linear. However, it is more common for clients to relapse and recycle though earlier stages. In the context of homework compliance, the six stages of change would suggest the following:

Stage 1. *Precontemplation.* At this stage, the client has not identified their specific contribution to their continued distress and have no intention of engaging with homework. When a client is in this stage of change, the therapist works to generate an exploration into considering the costs and benefits of change.

Stage 2. *Contemplation.* At this stage, the client realises that change is needed and is considering engaging with homework, but is unsure whether change would be worth the cost in time.

Stage 3. *Preparation.* The decision to change has been made by the client and the priority is identifying the focus of change and the strategies required in preparing to carry out homework.

Stage 4. *Action.* This stage represents active engagement with the homework process.

Stage 5. *Maintenance.* At this stage, relapse prevention strategies are considered and discussed.

Stage 6. *Termination.* At this stage, the client is ready to leave therapy and has established self-efficacy in the continued application of learned strategies.

Prochaska, DiClemente and Norcross (1992) emphasise that targeting interventions that match with the client's stage of change is important if change is to occur. With homework assignments, this has direct implications, as homework that does not consider the client's readiness for change may result in noncompliance with

the task. Furthermore, in selecting tasks relevant to the client's current stage of change, it is likely to aid in the effectiveness and compliance with the both the task and the therapy itself (Prochaska, Velicer, Rossi, Goldstein, Marcus, Rakowski, et al., 1994)

Protection Motivation Theory and Volitional Strategies

A number of social cognitive models of behaviour have been applied in developing health education interventions and have relevance in the present context. Protection motivation theory (Rogers, 1975, 1983) has been shown to be useful in both the intervention with and the prediction of health-related behaviours. This model proposes that intention to perform recommended health behaviour hinges upon two variables: threat appraisal and coping appraisal (S. Milne, Orbell, & Sheeran, 2000). The evaluation of fear relevant to one's endangerment from a threat disease/behaviour is known as threat appraisal. There are three mechanisms involved: perceived vulnerability, perceived severity and threat arousal. When there is a high perception of threat, the likelihood of adopting recommended health behaviour is also high. Coping appraisal is the evaluation of response efficacy, response cost and self-efficacy. When an individual believes a response will be effective, be low in cost and achievable, the intention to perform the behaviour will be more likely (Floyd, Prentice-Dunn, & Rogers, 2000; S. Milne, Orbell, & Sheeran, 2000).

However, Gollwitzer (1993) and Heckhausen (1991) suggest that motivation is solely the starting point of behavioural performance. They propose a model of action whereby the adoption of behaviour undergoes two distinct stages. The first stage is the motivational or deliberative phase, which equates to Protection Motivation Theory's intention formation where the benefits and costs of adopting behaviour are weighed

up. Within the homework compliance context, this equates to the client deciding whether or not the benefits (change) outweighs the costs (pain or distress).

Following this is a volitional phase (S. Milne, Orbell, & Sheeran, 2000) during which strategies and plans for the implementation are developed in order to ensure the behaviour is enacted. Putting this into the therapy environment, the client enacts strategies that allow homework to be performed.

Theory of Planned Behaviour

Central to the theory of planned behaviour (Ajzen, 1991) is the belief that the performance of any behaviour is determined by the interaction of behavioural intention and perceived behavioural control. Behavioural intentions summarise an individual's motivation to engage in behaviour; the more motivated people are to engage in the behaviour, the more likely they are to succeed in performing that behaviour. Perceived behavioural control reflects people's confidence in their ability to carry out a particular behaviour, and is regarded as being synonymous with Bandura's (1997) self-efficacy construct (e.g., Ajzen, 1991).

The extent of behaviour enactment is determined by the extent to which it reflects people's actual control over the behaviour, or their perceived behavioural control (Sheeran, Trafimow, & Armitage, 2003). In turn, behavioural intentions are guided by perceived behavioural control, subjective norms (perceived social pressure regarding behaviour performance) and attitude towards the behaviour (positive and negative evaluations of behaviour). Thus, motivation to engage in behaviour relies upon the behaviour having a positive appraisal, the individual feels social pressure to perform the behaviour, and believes they will be successful (Armitage, 2005).

Relating the theory of planned behaviour to homework compliance, a client's motivation, or behavioural intention, to engage with homework tasks will be incumbent on their perceived behavioural control and their attitudes and beliefs towards the homework task. Therefore, it is necessary, when assigning homework that these attitudes and beliefs about the homework ascertained in order to enhance homework compliance (A. T. Beck et al, 1979).

Measuring Compliance

The therapeutic value of homework has been demonstrated in a number of studies, with homework completion being positively related to treatment outcome (e.g. Addis & Jacobson, 2000). Similarly, a meta-analysis of 32 studies into the relationship between homework completion and outcome produced an effect size of .36 (medium effect) whilst highlighting that the focus has generally been on the quantity of homework completion, rather than on the quality of completion (Kazantzis, Deane & Ronan, 2000).

Springer and Reddy (2004) suggest a number of reasons why the assessment of client compliance with homework tasks is important. Where there is no clinical improvement, identifying whether or not a client is actually doing their homework would be significant especially when treatment efficacy becomes an issue. Similarly, as an aid in maintaining motivation, the ongoing monitoring and evaluation of homework assignments could have a positive impact. As well, when studying behaviour change, the monitoring of homework may help in identifying mechanisms involved in behaviour change (see also Doss, 1999).

The comparison of published research involving homework completion is hampered by the use of different methods of assessment. In a review of the assessment

of homework compliance, Kazantzis, Deane and Ronan (2004) determined that 72 percent of studies relied solely upon a single source of data such as the client, therapist or independent source. Of the studies reporting multiple sources of compliance data, two relied upon retrospective accounts utilising single item scales. Four studies were found to include both client and therapist data. In all, only four studies used the same measure of homework compliance, the Assignment Compliance Rating Scale (ACRS). Originally developed by Primakoff, Epstein and Covi (1986), the ACRS is a one item measure using a six-point likert scale with 1= *did not attempt any homework*, to 6= *the client did more than the assigned homework.*

Homework Rating Scale-II

Although the call for more encompassing measures of homework completion has been made (Helbig & Fehm, 2004), there exists no measure in current practice that comprises of both quantity and quality of homework completion. The recently developed Homework Rating Scale (HRS-II; Kazantzis, Deane & Ronan, 2005) addresses both quantity and quality of client homework completion as well as reviewing key aspects relating to the increased likelihood of engagement with homework tasks.

The HRS-II integrates both cognitive and behavioural theories in its construction (Kazantzis, Deane & Ronan, 2005). There are five main theoretically driven components of homework completion depicted in its 12 items: beliefs, situation, behaviour, consequences and synthesis. Conceptualised as a circular model beginning with the *designing* and *assigning* of homework phase, client beliefs about the homework affect whether or not they engage with the task. Situational factors may serve as potential barriers prior to the *doing* phase. In the doing phase, behaviour and

its consequences occur. Finally, during homework *review* phase, the experience is synthesised. Figure 1 depicts the integrated model of homework completion (Kazantzis, Deane & Ronan, 2005)

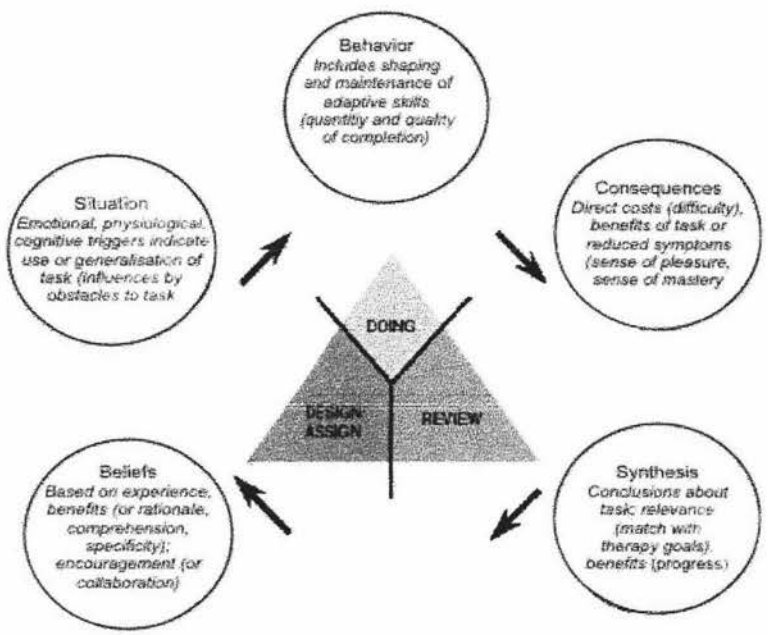

Figure 1. Integrated cognitive and behavioural model of homework completion. © 2005 by Nikolaos Kazantzis, Frank Deane, and Kevin Ronan. From the book "Using Homework Assignments in Cognitive Behavior Therapy", by N. Kazantzis, F. P. Deane, K. R. Ronan, & L. L'Abate (2005). New York: Routledge. Reprinted with permission.

Each of the five theoretical components of homework completion provide justification for the 12 items included in the HRS-II (Kazantzis, Deane & Ronan, 2005). As discussed in Chapter 2, cognitive theory posits that a client may have beliefs, attitudes or rules about homework that may influence the degree to which they comply with homework tasks. Therefore it is important to explore these beliefs. Four of the items on the HRS-II have been devised to assess the client's beliefs about the homework task itself as well as the process in designing and assigning the homework. The HRS-II rates the extent to which the client understood the reason for the homework task (*rationale*), found the guidelines for carrying out the task specific (*specificity*), played and active part in the planning of the homework (*collaboration*) and the degree to which they understood what to do (*comprehension*).

As noted by A. T. Beck et al (1979), clients may fail to complete, or even begin, homework tasks due to negative attitudes towards the homework. As described in Chapter 2, negative attitudes and beliefs are developed through prior experience and may appear when life events match these beliefs. Further, As well, there may be practical obstacles that may prevent a client from undertaking homework, such as scheduling difficulties due to family responsibilities. In order to assess client noncompletion, there is one item addressing the extent to which the client experienced barriers to completing the homework (obstacles).

The degree to which the client carried out homework assignments is addressed by examining how much homework was completed (*quantity*) and the level of learning gained through undertaking the homework (*quality*). These aspects are important to assess, as client variation in homework completion occurs (Kazantzis, Deane & Ronan, 2004). The splitting of the construct into two components allows for a more definitive assessment of overall completion that does not confound quantity with

quality, which may or may not be the case with previous measures of homework completion.

As discussed in Chapter 2, operant conditioning proposes that there is a link between behaviour and its consequences. Similarly, cognitive theory posits that the perceived cost-benefit ratio of homework assignments is influenced by intrinsic, or internal, consequences resulting from doing the homework. To this end, three items on the HRS-II examine the interpretation of consequences resulting from doing the homework. Specifically, the HRS-II evaluates whether or not the client interpreted the task as difficult (*difficulty*), found the activity enjoyable (*pleasure*) and found a sense of control over their problems consequent to completing the task (*mastery*).

The review of homework allows for the therapist to facilitate client synthesis of the homework process and to make conclusions derived from the learning involved in doing the homework. Specifically, synthesis of the client's experience is ascertained based on ratings of how well the homework matched goals for therapy (*match with therapy goals*) and whether or not the homework task was found to be helpful (*progress*).

The Integration of Homework into Practice

While client motivation to engage in homework depends on a number of possible factors, the skill of the therapist plays an important role in promoting and maintaining client engagement with out of session tasks. Homework in CBT is an integral part of therapy (e.g., A. T. Beck et al., 1979; J. S. Beck, 1995; Detwieler & Whisman, 1999; Persons, Davidson, & Tompkins, 2001, Scheel, Hanson, & Razzhavaikina, 2004; Thase & Callan, 2006). Despite the attention that homework has received in research, it has only been recently that the theoretical and empirical

foundations for homework integration in CBT have been distilled into models for practice.

Early recommendations for the integration of homework into practice grew from the increase in the use of behavioural formulations to describe behavioural difficulties (e.g., Kanfer & Phillips, 1966). This led to the development of a clinician's guide consisting of 150 homework assignments that therapists could use to treat specific behavioural problems (Shelton & Ackerman, 1974). Within this manual, guidance was provided as to how these tasks could be implemented by the therapist in their practice.

One of the first specific "model of practice" outlining recommendations for homework integration into behavioural therapy practice was devised by Shelton and Levy (1981). This model centres on the assignment of homework, specifically on the how, when and how often homework tasks are to be implemented. The Shelton and Levy (1981) model was similar to recommendations previously made by A. T. Beck et al. (1979); the recommendation that each session should begin and end with a discussion about homework and that homework be assigned with a great deal of behavioural specificity.

However, there has been some criticism of the Shelton and Levy (1981) model. Criticism centres on the inadequate flexibility in accommodating the range of homework in CBT and that the role of the therapeutic relationship in homework completion is not sufficiently addressed by the model (Kazantzis, MacEwen, & Deane, 2005). Further, while practitioner surveys report the widespread use of homework, although it is reported that practitioners often do not comply with these recommendations (Kazantzis & Deane, 1999; Kazantzis, Busch, Merrick, & Ronan, 2004; Kazantzis, Ronan, & Deane, 2005).

Detweiler and Wishman (1999) propose a heuristic model integrating therapist, client and task factors, thus presenting a much more balanced model than those previously proposed that focused solely on therapist behaviours (e.g., Shelton & Levy, 1981). Although derived through empirical findings, the Detweiler and Whisman (1999) model is not theoretically derived. Further, the focus of the model is on client compliance on client homework compliance, and is not considered a guiding model for practice (Kazantzis, MacEwen, & Datillio, 2005).

Recently, Scheel, Hanson and Razzhavaikina.(2004) propose a model for homework integration that is based on both empirical and theoretical grounds. Drawing from findings and recommendations made by Kazantzis and Deane (1999), the model conceptualises homework integration as a six-phase model: client-therapist formulation, therapist delivery, client receipt, implementation, therapist asking about homework compliance and client report of homework experience. Each stage is accompanied with proposed practice strategies, with each justified by empirical support (Scheel, Hanson, & Razzhavailina, 2004). However, the theoretical foundations of homework has not been incorporated into the practice strategies of the model, and as such, only provides a limited guiding model for practice.

Guiding Model for Practice

Kazantzis, MacEwen and Datillio (2005) have synthesised existing recommendations into three distinct stages of homework integration: homework design, homework assignment and homework review. This guiding model for practice presents a manualised procedure aiding the process of homework integration into CBT sessions (Kazantzis, MacEwewn & Dattillio, 2005). Further, the guiding model consolidates and enhances previous recommendations for homework integration, and

bases the new recommendations securely on theoretical and empirical foundations. Drawing on theoretical and empirical foundations of homework, the authors also incorporated clinical recommendations made by contributing authors to the book. The resulting guiding model is a manualised protocol for future prospective process-outcome research.

The authors (Kazantzis, MacEwewn & Datillio, 2005) propose three advantages over previous recommendations for practice: (a) a focus on the facilitative qualities of the therapeutic relationship, therapist qualities and therapist beliefs; (b) the grounding of recommendations on cognitive and behavioural foundations; and (c) the individualised tailoring of both the content and the process of homework.

The model, shown in Figure 2, conceptualises homework integration as a cyclical, three-stage process involving the review of previously assigned homework, the design of new homework and the assignment of the homework (Kazantzis, MacEwen & Datillio, 2005). Each stage is comprised of a number of criterion behaviours associated with that stage. The following sections will briefly detail the key therapist behaviours (shown in italics) identified by the authors within each stage of homework integration.

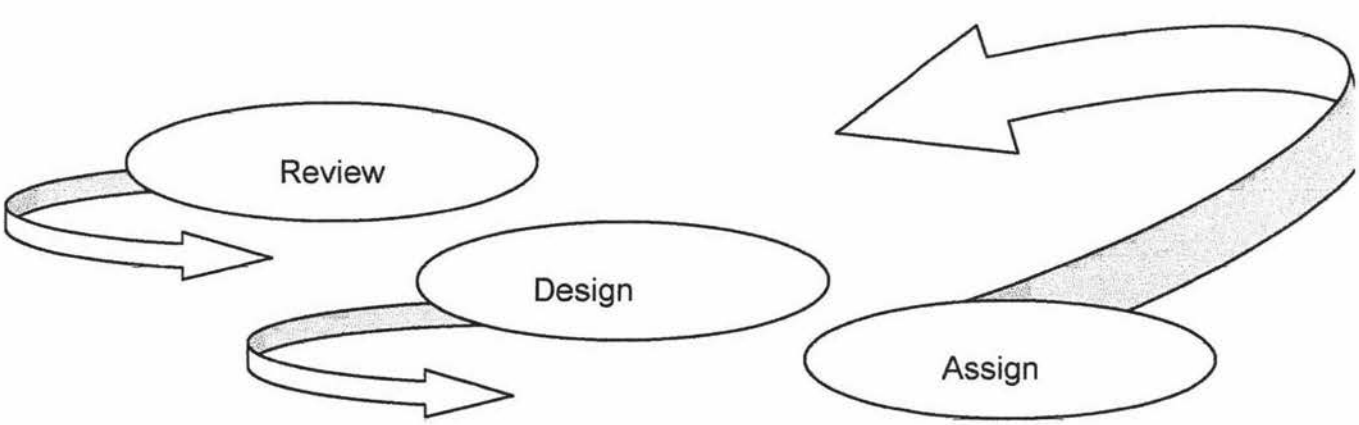

Figure 2. Cyclical process for recommending homework (Kazantzis, MacEwen & Datillio, 2005). Used with permission.

Homework Review

The *review* of homework is a crucial part of a session in that it reinforces the belief that compliance is important in promoting improved outcome (Worthington, 1986). Similarly, Bryant, Simons, and Thase (1999) found that the review of previously assigned homework was predictive of clients' subsequent compliance with homework. The provision of therapist *praise* for client homework completion is one component of the contingencies for homework completion, with others including intrinsic contingencies (such as reduced distress).

Whereas most research on homework compliance has examined the quantity of homework completed, the few studies that have looked at the degree of learning gained through skill acquisition have found that *quality* is a better predictor of outcome than quantity of homework completion (e.g., Schmidt & Woolaway-Bickel, 2000). Following on from quality, therapists should endeavour to obtain a *conceptualisation of a specific situation* during which homework was completed (Persons), thus gathering information concerning the client's beliefs about the homework. In addition, the review of homework should identify any barriers that the client may have encountered in doing their homework and provide an *individualised conceptualisation for homework non-completion* (Kazantzis, MacEwen & Datillio, 2005). Should there be any practical obstacles that prevented homework completion, the therapist should work with the client to problem-solve these, as this information may be important to consider in the design of future homework tasks.

Homework Design

In designing homework tasks, the individualised case *conceptualisation* should be taken into account (Beck et al., 1979) as the choice of a homework task that is not

relevant to the client's problem may result in homework noncompliance (e.g. Friedberg & McClure, 2005; Pachana & Sofronoff, 2005) and decreased task acceptability for the client (Conolet, Padula, Payton, & Daniels, 1994). Along with this, the provision of a clear *rationale* linking the particular homework task with treatment goals serves to reinforce the collaborative nature of CBT A. T. Beck et al, 1979). Further, *collaboration* in task selection, ensures that tasks are guided by the client's goals and priorities (Young & Beck, 1982) Socratic questioning and *guided discovery* often prove useful in identifying and client coping strategies and beliefs about the homework, as well as determining the *client's ability and perception of the difficulty of the task*, thus problem-solving potential barriers to compliance (Leahy, 2002; Tompkins, 2004). Similarly, the therapist can utilise a situational conceptualisation to identify beliefs and situational triggers that may impact on homework completion (Kazantzis, MacEwem & Datillio, 2005).

The designing of homework tasks also includes the use of *in-session practice*, which serves as a bridge between rational and experiential learning. Modeling of behaviour by the therapist, or the actual in-session practice by the client helps to ensure that the client has some experience prior to undertaking homework tasks on their own. *Guided imagery* used in task rehearsal can be useful in anticipating both predictable and unpredictable outcomes related to task engagement (Stokes & Baer, 1977).

Homework Assign

The assignment of homework needs to include *specifics* with regards to how, when, where, with whom and duration in order to limit misunderstanding by the client (Shelton & Levy, 1981; Kazantzis & Deane, 1999). Further, misunderstanding can be reduced through provision of a client summary of the rationale for the

homework (Kazantzis, MacEwen, & Datillio, 2005). The assigning of homework must also assess practical obstacles and potential difficulties in homework completion and stress the learning experiment focus of the homework (A. T. Beck, et al.,1979; Pesrons, 1989). Finally, the client should summarise the homework and the therapist obtain an indication of client readiness for homework engafement.

While the use of between-session tasks is widespread across different therapeutic approaches, there has been little empirical research into factors promoting homework completion (Detweiler & Whisman, 1999; Scheel, Hanson & Razzhavaikina, 2004).

As part of the Cognitive Behavior Therapy Project, a manualised protocol for effective homework integration was developed (Kazantzis, MacEwen, & Dattilio, 2005). Within this protocol is the guding model for practice. While derived from other models for homework integration, the manualised procedure has a number of advantages over prior recommendations for practice. Firstly, the authors Kazantzis, MacEwen and Dattilio (2005) have developed a "*guiding model for practice*" based on both theoretical and empirical foundations. This model has formed the basis of a theoretically driven measure of therapist adherence and competence in the administration of homework assignments in CBT (Homework Adherence and Competence Scale: HAACS; Kazantzis, Wedge & Dobson, 2005).

Although recommendations for a systematic approach towards homework integration into CBT exist (Beck, et al., 1979; Shelton & Levy, 1981), the actual practice of incorporating homework within clinical practice does not necessarily take such a structured approach (Kazantzis & Deane, 1999). Whereas some aspects of therapy have been greatly researched, the manner in which homework is integrated into CBT has not yet been systematically studied.

Summary

Client motivation to complete homework depends on any number of factors. While homework is often studied, there have been few studies using the same measure. Further, the assessment has tended to be on quantity rather than quality. In order to address limitations in homework assessment, Kazantzis, Deane & Ronan (2005) have developed a new measure of homework completion that encompasses both the cognitive and behavioural determinants of homework completion. As therapist behaviours can impact on client homework completion, this chapter also described recommendations for the integration of homework into practice, from early recommendations through to the recently manualised homework protocol developed by Kazantzis, MacEwen, and Dattilio (2005).

CHAPTER 4

Therapist Competence and Adherence

Overview

The accurate assessment of treatment integrity is dependant on the use of measures specifically designed to assess both adherence and competence (Perepletchikova & Kazdin,2005). In a review of the literature on the effects of manual-based training on treatment fidelity, Miller and Binder (2002) concluded that there is little empirical support for collapsing adherence and competence into one variable, nor should adherence be considered a tenable substitution for competence. Although Waltz et al. (1993) estimated that the cost of assessing competence was 10 times that of adherence assessment, they found that adherence was not significantly associated to overall competence ratings made by expert raters ($r = .34$). Further, Barber and Crits-Critsoff (1996) suggest that relying on adherence alone may be confounded by raters inadvertently assessing competence rather than adherence.

This chapter will outline the definitional issues related to therapist adherence and competence relative to the assessment of treatment integrity in randomised clinical trials. Further, the importance of assessing therapist adherence and competence will be discussed, as will suggestions from research regarding the assessment of treatment integrity. Research findings concerning therapist competence and adherence in integrating homework will be considered, with specific regard to assessment measures used. Finally, the development and utility of the Homework Adherence and Competence Scale (HAACS; Kazantzis, Wedge, & Deane, 2005) will be briefly outlined.

The Importance of Treatment Integrity

Inferences concerning treatment effects require some form a guarantee that the treatment was provided as intended, or designed (Moncher & Prinz, 1991;Vermilyea, et al., 1984). Treatment integrity is akin to quality control in that it enables researchers and clinicians to be better able to say that a specific treatment effected change rather than change due to other factors not accounted for. The ability to assess treatment integrity is especially important in treatment outcome research, as the external and internal validity of a study rests upon the assumption of both therapist adherence and competence in administering therapy (Perepletchikova & Kazdin, 2005; Moncher & Prinz, 1991).

As Gresham (2005) points out, treatment integrity is necessary in determining whether or not there is a functional relationship between therapeutic procedures and treatment outcome. There are, however, differing views on whether or not there is a relationship between treatment integrity and therapeutic change. For example, treatment adherence was not associated with treatment outcome for both the cognitive behavioural and interpersonal psychotherapy conditions in the National Institute of Mental Health Treatment for Depression Collaborative Research Program (NIMH TDCRP; Elkin, 1999) even though patients in both conditions showed significant reduction of depressive symptoms. Conversely, in a large, multi-site outcome study of a manualised Interpersonal Psychotherapy protocol, there was a moderate relationship found between therapist competence and patient assessment of overall outcome (r = .56). However, the competence ratings were based on client reports at the end of the study and therefore, these ratings may indicate the phenomenon of the "grateful testimonial", that is, those in receipt of a service tend to

provide generous ratings for that service as is suggested by D. Milne (1987). In a recent report on therapist effects in the NIMH TDCRP, Elkin, Falconnier, Martinovich and Mahoney (2006) proposes that the impact of therapist effect on outcome may best be addressed within the context of naturalistic studies. Crits-Critsoff and Gallop (2006) suggest that therapist effects may be, at best, between small to moderate.

So why is it that there is no clear association between treatment integrity and treatment outcome? It may be that the measure of treatment integrity used does not adequately measure treatment fidelity, as is suggested by Shaw et al, (1999). Another possibility is the reliance on posttreatment integrity data, which precludes the identification and rectification of therapist drift throughout the course of therapy (Gresham, 2005). Further, client difficulty may impact on therapist performance, producing inconsistencies in treatment delivery (e.g., Foley, O'Malley, Rounsaville, Prusoff, & Weissman, 1987).

Therapist Adherence and Competence

Within randomised clinical trials, traditional integrity checks involve the random sampling of between 20 and 30 percent of sessions in order to check for adherence to treatment protocols (Wampold & Bhati, 2004). While adherence refers to the degree to which a therapist 'follows' the treatment manual, competence refers to the degree to which a therapist 'effectively' conducts the treatment (Nezu, A. M. & Nezu, C. M., 2005). Although related, adherence does not equate with competence (Elkin, 1999; McGlinchey & Dobson, 2003; Wampold & Bhati, 2004). However, adherence is an essential condition for competence (Waltz, Addis, Koerner, &

Jacobson, 1993). Dobson and Singer (2005) describe a hierarchical relationship between competence and adherence. Therefore, it is recommended that the assessment of both need to be considered in determining treatment integrity.

Further, it has been noted that little is known about the temporal course and stability of therapist adherence and competence across therapy sessions (Svartberg, 1999). There is some indication that therapist competence maintains a moderate degree of temporal stability (Barber, Chrits-Christoff, & Luborsky, 1996). Svartberg (1999), employing hierarchical linear modeling, found little variation in therapist competence

As a component of treatment integrity, adherence is the answer to following question: Did the client receive only what was prescribed by the treatment and not what was proscibed? In effect, adherence is like a checklist determining how much of the treatment protocol was provided. Competence is related in that it is the skill with which the prescribed elements are delivered.

Assessment of Therapist Adherence and Competence

In order to be able to assess the extent of treatment fidelity, the components of treatment need to be identified and detailed in research protocols or treatment manuals (Dobson & Singer, 2005; McGlinchy, et al., 2003) As suggested by Perepletchikova and Kazdin (2005), the detailing of necessary components is a necessary step in creating measures of treament integrity. In effect, the identification of specific therapist behaviours provides the framework for adherence, whereas competence is derived by the assessment of the quality of that delivery.

In research, therapist factors such as competence and adherence have generally been treated as non-random, or fixed variables, which can be controlled for statistically. However, Kazdin (2003) suggests that therapist factors such as competence and adherence to treatment protocols impact directly on the validity of randomised clinical trials. This said, therapist variability, or effects, in clinical trials has not generally been found to be statistically significant (Crits-Cristoff et al., 1991). Elkin (1999) reports that in the original National Institute of Mental Health (NIMH) Treatment of Depression Collaborative Research Program (TDRP) study, preliminary analyses of the effect of therapist variability on outcome did not reach statistical significance, therefore therapist variables were not included in the major analyses. However, she states that the lack of significant findings "seemed due not so much to small mean differences among the therapists, as to the large variability within therapists" (Elkin, 1999, p. 13). In fact, Crits-Critsoff and Mintz (1991) recommends using p values greater than .2 or .3 in lieu of the standard .05 and that when variability between therapists may exist, the therapist should be treated as a random variable in order to decrease the risk of Type I errors. That is, not finding differences when there are, in fact differences.

Assessment of Therapist Competence and adherence in CBT

The move towards the identification of empirically supported treatments and the push towards the standardisation of therapy through manualisation (see Addis & Krosnow, 2000; Chambless et al., 1996; Waltz et al., 1993) has led to the development of purpose-built measures of treatment integrity. Similarly, training institutions are increasingly reliant on the use of these measures to ensure the quality of the trainee

therapists (Lau, Dubord, & Prikh, 2004; D. Milne, Woodward, Hanner, Iceton, Fitzsimmons, & Rochester 2003; Sudak, J. S. Beck, & Wright, 2003; Reichelt, James, & Blackburn, 2003; Weerasekera, et al., 2003). CBT is a fairly structured therapy, it is not surprising that there are a number of measures that have been designed to assess therapist adherence and competence in administering CBT (Barber, Liese, & Abrams, 2003; Dobson & Kazantsis, 2003). Among these are the Cognitive Therapy Scale (CTS; Young & Beck, 1980), the Collaborative Study Psychotherapy Rating Scale (CSPRS; Elkin, et al. (1989).) and the Cognitive Therapy Adherence and Competence Scale (CTACS; Barber, Liese, & Abrams, 2003).

While these measures have been designed to specifically assess treatment integrity in CBT, or include a subscale measure for CBT in the case of the CSPRS, there have been a number of concerns regarding the overall utility of the measures (Kazantzis, 2003; McGlinchey & Dobson, 2003). Perhaps the most telling is that both the CTS and CSPRS do not measure both competence and adherence (Barber, Liese, & Abrams, 2003). Although the CTS, CSPRS and CTACS provide a range of therapist behaviours associated with CBT, none have been widely validated, with the CTACS receiving validated within CBT for substance abuse (Barber, Liese, & Abrams, 2003; Kazantsis, 2003).

Therapist Competence and Adherence in Homework Integration

Although the role of homework in psychotherapy has been receiving increasing attention within CBT (Detweiler & Whisman, 1999; Scheel, Hanson, & Razzhavaikina, 2004) the assessment of its delivery has generally been limited to one or two questions on a larger measure of overall competence such as the case with the

CTS (Young & Beck, 1988). One notable exception is the recently developed Therapist Homework Assignment Competency Scale (THACS; Bryant, Simons, & Thase, 1999), which was designed to separately assess both the adherence and the competence in the administration of homework. However, the four-item measure was not developed based on theory, but rather, based on the CTS. Even so, initial data support the review of homework as a predictor of treatment outcome (Bryant, Simons, & Thase,1999). However, Startup and Edmonds (1994) failed to find a relationship between therapist adherence to homework procedures and outcome, when adherence was measured by therapist report. Detweiler-Bedell and Whisman (2005) found two therapist behaviours were associated with improved outcome: the discussion of barriers to homework compliance among less involved clients and the setting of concrete goals.

Homework Adherence and Competence Scale

Within CBT, the recommended techniques for homework administration include providing a rationale for the assignment, being collaborative, reviewing the homework in subsequent sessions, and being specific about the task itself (A. T. Beck, et al., 1979). As outlined in the Chapter 3, a number of recommendations have been put forth regarding the integration of homework into clinical practice (see Detweiler & Whisman, 1999; Kazantzis & Deane, 1999; Scheel, Hanson, & Razzhavaikina, 2004). Kazantzis, MacEwen and Dattilio (2005) have developed a "guiding model for practice" for homework integration that is based on both theoretical and empirical foundations. This model has formed the basis of a theoretically driven measure of therapist adherence and competence in the administration of homework assignments

in CBT (Homework Adherence and Competence Scale: HAACS; Kazantzis, Wedge & Dobson, 2005). A main point of difference between this measure of therapist competence and adherence in homework integration is the separation of the two constructs. The utility of this is the potential to provide separate scores for adherence *and* competence, a suggestion which has been proposed by recently by Perepletchikova and Kazdin, (2005). Further, this scale differs from other measures on the derivation of competence ratings; competence ratings are possible *only* for items adhered to by the therapist, thus, maintaining the hierarchical nature of treatment integrity.

The HAACS (Kazantsis, Wedge & Dobson, 2005) is an observer rated measure consisting of three sections: homework review, homework design and homework assignment. The 19 items reflect key therapist behaviours from Kazantzis, MacEwen and Dattilio's (2005) *Guiding Model for Practice*. Two of the originally recommended therapist behaviours (e.g., recording homework completion in session notes and making a written note of the assigned homework) were not included in this measure, due to the difficulty in assessing the degree of competence for these behaviours when rated by observation.

The measure includes both adherence and competence ratings for each item. Adherence is initially scored as a dichotomous yes-no, whereas competence is measured on a seven-point likert scale, with 0 equated with non-adherence and 6 signifying excellence in administration. The anchors were developed explicitly for each of the nineteen items in order to provide behaviours specific to that item. For a more detailed account regarding the development of the anchors, see Wedge (2005).

Each section has an associated global rating, thus providing observers with the opportunity to provide an overall competence rating for each section. The global

ratings allows for flexibility in scoring reflecting any special contextual factors (such as a difficult client) or considerations of the appropriateness of not adhering to certain items.

Two pilot studies of the HAACS (Wedge, 2005) indicates an excellent level (Cichetti, 1994; Cichetti & Sparrow, 1981; Fleiss, 1981) of overall rater agreement, for adherence items using the percentage agreement method for the estimation of reliability (range 74% to 85%). Percentage agreement for individual items ranged from 40% to 100%, which suggests that some items may have been problematic for the raters. However, the raters consisted of four graduate clinical psychology students and one first year registered clinical psychologist, and the relative inexperience of the raters in CBT may have impacted on the ratings of some items. Interrater reliability for overall competence ratings using intraclass correlation coefficients, ranged from .79 to .83, indicating an excellent degree of reliability between raters (Cichetti, 1994; Cichetti & Sparrow, 1981; Fleiss, 1981).

The Therapist Quick Reference (TQR; Figure 3; Kazantzis, Deane, & Ronan, 2005), used in conjunction with the HAACS, outlines the key therapist behaviours identified in the *Guiding Model for Practice* in Chapter 3. Each behaviour lists the corresponding item number used in the HAACS.

Summary

In the field of psychotherapy outcome research, it is important to ascertain therapist adherence and competence to treatment protocols in order for there to be any degree of internal and external validity associated with the study. Further, the issue of therapist competence and adherence is not limited to research, but is an

essential part of certification and ongoing training and supervision (D. Milne, & James, 2002; Pretorius, 2006; Reichelt, James & Blackburn, 2003). Further, this assessment must be comprehensive and not be limited to adherence alone (Miller & Binder, 2002; Perepletchikiva & Kazdin, 2005). The assessment of therapist competence and adherence in homework integration has routinely been assessed by single item measures. The recent development of a guiding model for practice in homework integration (Kazantzis, MacEwen & Datillio, 2005), has resulted in the creation of a comprehensive measure of therapist competence and adherence, the HAACS (Kazantzis, Wedge & Deane, 2005).

Therapist's Quick Reference* - Adapted for Rating the HAACS

REVIEW	DESIGN	ASSIGN
1. Discuss completion of previously assigned homework.	**6.** Discuss new or revised homework during session.	**15.** Summarize homework rationale in relation to therapy goals.
2. Provide verbal reinforcement for any portion carried-out.	**7.** Use guided discovery to identify coping strategies and beliefs.	**16.** Specify how the task will be practically possible (i.e., when, where, how often, and how long it will take).
3. Use a situational conceptualization to review previously assigned homework.	**8.** Integrate a disorder specific cognitive model and individualized conceptualization.	**17.** Consider potential difficulties.
4. Use an individualized conceptualization to make sense of any portion of non-completed homework.	**9.** Collaboratively select task.	**18.** Emphasize learning 'experiment' focus.
5. Problem solve practical obstacles.	**10.** Present a rationale that aligns with the client's treatment goals.	**19.** Summarize the homework and obtain indication/ratings of readiness, importance, and confidence
	11. Ask about client's ability and perceived task difficulty.	
	12. In-session practice of task.	
	13. Use guided imagery to begin experiential learning in-session.	
	14. Use a situational conceptualization to identify beliefs and situational triggers.	

age number

Figure 3. Therapist's Quick Reference* - Adapted for Rating the HAACS

CHAPTER 5

Psychometric Theory

Overview

The choice of specific assessment measures or tools in both empirical research and clinical practice is dependent upon the dimensions of interest, be it symptoms, behaviours or cognitions. Often, there are specific measures available from which to choose. There are, however, requirements that must be met to ensure the applicability and utility of the measure based upon its intended purpose (Nunnally & Bernstein, 1994). More specifically, the validity and reliability evidence of any potential measure to be utilised in empirical research or clinical practice must be assessed critically in order to ensure the soundness of any inferences drawn from the data produced (Kline, 1998; Messick, 1995).

This chapter will look at the importance of psychometric theory in choosing psychological assessment measures, as well as the process by which evidence supporting both the validity and the reliability of a test's scores is accrued. Further, existing guidelines surrounding the reporting of test properties and results will be outlined along with results from recent analyses of the reporting of psychometric evidence in psychology and counselling journals. Finally, of a brief overview of psychology journal articles published in 2005 involving the psychometric evaluations of new, or adapted measures in the field of clinical psychology will be included to show the current state of reporting practices,

Psychometric Theory

Psychometrics is the science of psychological measurement and is grounded in classical test theory, which posits that within observed scores there is both the "true score" plus error (Cronbach, 1971; Kline, 1998). Error can be either systematic, where non-random error reflects a possible systematic bias, or random, where error is due to random traits of the subjects (Kline, 1998). The ratio of the true score to the observed score (true score + error) reflects the reliability of the observed score. The inherent existence of error within observed test scores requires the systematic evaluation of measurement instruments in order to minimise the intrinsic error and its effect upon the validity and reliability of the data produced (Nunnaly & Bernstein, 1994). Further, the existence of error impacts upon the inferences that can be drawn from the data (Kline, 1998).

The role of psychometrics is to minimise the inherent error in test scores, thus producing scores closer to the true score. This however, as Kline (1998) suggests, is dependent upon the true score actually being representative of the construct being assessed. Therefore, an observed score may be reliable (high true score, low measurement error), yet may not be valid. Reliability and validity are the two most important aspects to consider when choosing an assessment measure (Nunnally & Bernstein, 1994), along with the understanding that the validation of a measure is an ongoing process, and according to some, should be an important aspect of any empirical research (Messick, 1995).

Measurement in Psychotherapy Research

Empirical research in the field of clinical psychology is generally concerned with two different aspects: psychotherapy process and psychotherapy outcome. The assessment of psychotherapy process is the focus on what occurs within the psychotherapy session and is differentiated from outcome research where the focus is on the assessment of either long-term or immediate changes due to the effects of therapy (Hill & Lambert, 2004).

The choice of appropriate assessment tools, or measures, in clinical psychology research should be guided by the psychometric properties of those measures (Kazdin, 2003). The reliability and validity evidence accrued though prior research can aid the researcher in their choice. However, as new areas of research focus are explored, the development of new measures is inevitable.

Unlike the natural sciences, where the object of interest usually exists in the 'real' world, the objects of interest in psychological research are generally not self-evident (Kline, 1998). The difficulty in measuring psychological constructs, necessitates the validation of all measures used in empirical research in psychology.

Measurement in psychotherapy research is reliant on the ongoing validation process of a measure's psychometric properties, especially when the construct of interest may be within an emergent area of interest. As research within the emergent field expands, the development of new measures designed to tap specific constructs of foci is a very real reality (Kazdin, 2003; Messick, 1995). In turn, the gathering of validity and reliability evidence of these new measures must be determined in order to ascertain the ultimate value of a study (Kazdin, 2003).

Process measures, such as therapist adherence and therapist competence measures have been developed in order ensure the reliability of findings from randomised clinical trials (Elkin, 1999). The overarching construct of interest regarding therapist competence and adherence is treatment integrity, or fidelity (McGlinchey & Dobson, 2003; Waltz, Addis, Koerner & Jacobson, 1993). The identification of deviations from therapy protocols, allows for researchers to maintain consistency in the delivery of psychotherapy (Elkin, 1999; McGlinchey & Dobson, 2003). Thus, treatment integrity is an area where accurate and ongoing measurement is necessary within empirical research (Elkin, 1999; McGlinchey & Dobson, 2003; Okiishi, Lambert, Nielsen & Ogles, 2003; Waltz, Addis, Koerner & Jacobson, 1993).

As with most aspects of empirical research, sample size can severely impact on the ability to form inferences regarding test scores (Cohen, 1992). The smaller the sample, the more likely that variation in test scores due to chance will impact upon correlation, leading to the artificial inflation of a validity coefficient (Kaplan & Saccuzzo, 2001). Further, small sample size can yield low reliability estimates (Thompson, 2003). The use of statistical power analysis in research can indicate the required sample size in order to detect effect sizes (Cohen, 1995). The reporting of effect sizes and confidence intervals are both recommended reporting practices (Wilkinson, & Task Force for Statistical Inference, 1999).

The Reporting of Psychometric Evidence

The validation process of measurement instruments inevitably is linked to statistical procedures designed to answer specific research questions (Kline, 1998; Nunnally & Bernstein, 1994). The current reference standards for the reporting of

statistical inference is outlined in the fifth edition of the American Psychological Association *Publication Manual* (APA, 2001). Similarly, the APA's Task Force on Statistical Inference (TFSI) has published recommendations regarding the use of statistical methods in research (Wilkinson & TFSI, 1999).

In assessing the adequacy of scales' psychometric properties, it is necessary to rely on the reporting of reliability and validity evidence (Meier & Davis, 1990). However, the rationale for using a specific measure rarely extends beyond the inclusion of a citation that, according to Meier and Davis (1990), "presumably contains quantitative and qualitative information" (p. 113). Wilkinson and TFSI (1999) recommend the inclusion of validity evidence as well as reliability coefficients of the scores, regardless of whether or not the focus of the research is psychometric.

Psychometric evidence is not a property of an assessment measure, *per se*, but refers to the reliability and validity of the scores and inferences yielded from the administration of a measure (Wilkinson & TFSI, 1999). Thus, the reporting of psychometric evidence is a necessary part of the ongoing process of scale validation. Although validity is generally regarded as fundamental in both the construction of, and evaluation of, measures, the reporting of psychometric properties of assessment measures in psychology journals reflects an underreporting of validity evidence (Hogan & Agnello, 2004). Recent reviews of reporting practices reveal that although reliability statistics are by far the most common psychometric evidence outlined in psychology journals (Hogan & Agnello, 2004; Meier & Davis, 1990; Qualls & Moss, 1996), there is still a need for improved reporting practices.

In summarising methodological trends in reporting psychometric properties in articles in the *Journal of Counseling Psychology* from 1967, 1977 and 1987, Meier and

Davis (1990) found an increase in reporting both reliability and validity evidence. However, by 1987, only 40% of articles provided some form of reliability estimate, and only 5% included any form of validity evidence. Similar findings regarding the reporting of reliability information (41%) were found by Qualls and Moss (1996) in summarising 2,167 measures used in 622 studies from 22 APA journals published in 1992. They did find differences in the reporting of validity, with 32% providing validity information. A surprising 49% of articles did not provide any reliability or validity information, and only 20% supplied both reliability and validity evidence.

Hogan and Agnello's assessment of reporting practices concerning measurement validity (2004) sampled 696 research reports listed in the APA's *Directory of Unpublished Mental Experimental Mental Measures* and found that only 54.6 % of reports included any type of validity evidence. Of theses, only 2.3% reported two types of validity evidence and none reported more than two types of validity evidence. Correlation with other variables was the predominant validity evidence provided (approximately 90%), with only 4% utilising group contrasts and 2% using factor analysis. There was no mention on content validity in any of the measures reviewed, a finding supported by Whittington (1998). Given that content validity affects the inferences that can be drawn from the obtained scores (Haynes, Richard, & Kubany, 1995), and that content validity may degrade over time due to evolving theories (Cronbach, 1971), it is surprising the paucity of content validity reported.

Based on the above reviews of reporting practices, it is evident that the evaluation process of new measures be as comprehensive as possible and should include multiple aspects of psychometric evidence, not simply validity evidence or estimates of reliability, but both. Further, the reporting of validity evidence should not

be limited to correlational evidence, but should include some evidence of content validity. As this project includes the gathering of psychometric evidence for two new measures, the requirements for sufficient supportive evidence provided specific strategies for the gathering of this evidence. Of particular importance was the gathering of content validity evidence, which was outlined in the preceding chapters. That is, the theoretical and empirical grounds upon which the HRS-II and HAACS were developed.

Validity Evidence

It is a requirement for researchers to provide validity evidence for all measures used in empirical research in order to demonstrate that the measures are in fact assessing what they are purported to assess (Hogan & Agnello, 2004; Messick, 1995). There are numerous types of validity evidence including face, construct, content, convergent, discriminant, predictive, factorial and consequential validity. However, the consensus appears to be that all categories of validity are subsumed by construct validity (see Messick, 1993, 1995).Messick (1995) and Cronbach (1971)describes validity as constantly evolving process. Validity, or the accuracy of measurement, is based upon inferences drawn from available evidence and is not measured directly. As stated previously, correlational methods appear to be the most common type of validity evidence from which validity is inferred (Hogan & Agnello, 2004). By this, a measure is shown to have validity evidence by way of correlating with other, existing criterion measures (Kaplan & Saccuzzo, 2001). As posited by Messick (1995), validity is a unitary construct, and as such, the accrual of validity evidence, through various inferences, lends more and more confidence that a measure is measuring what it

purports to measure. However, it must be recognised that validity evidence only applies to the inferences derived from previous research (Cronbach, 1995). The more a measure is used, the more validity evidence is acquired. Validity is not a static property, but is dynamic, and as such, validation of a measure must be approached as an evolving and ongoing process (Cronbach, 1971; Messick, 1995).

The notion of test validity in clinical psychology research, extends to the concept of clinical sensitivity of the measure. That is, how well a measure is sensitive to change due to treatment effects (Jacobson, et al., 1999; Kendall, Marrs-Garcia, Nath, & Sheldrick,.,1999; Taylor,1995). Taylor (1995) suggests that effect size for differences is a good indicator of a measure's sensitivity to treatment effects, with observer-rated scales producing larger effect sizes than self-report scales. However, Jacobson et al. (1999) posits that effect size alone is insufficient as an indicator of clinically significant change, and Kendall, Marrs-Garcia, Nath, and Sheldrick (1999) recommends the use of normative comparisons in providing evidence for clinically significant change. While there are differences in what evidence is sufficient in determining the clinical sensitivity of a measure, all require comparisons based on the mean scores between two or more groups. Further, none suggest the abandonment of the statistical procedures used in psychometric evaluations, but recommend the determination of sensitivity in conjunction with statistical reporting procedures.

Reliability Evidence

As outlined earlier, reliability is the ratio of the true score to the total, or observed, score. Again, the observed score contains both true score and measurement error. Reliability relates to the extent to which an item, measure, or scale will produce

the same score upon subsequent administrations, populations or locations providing that there are no differences in relevant variables in the administrations (Kline, 1998). Statistically speaking, reliability coefficients are generally produced through correlational procedures.

Unlike validity evidence, which cannot be measured directly, the reliability of obtained scores can be computed directly (Kaplan & Saccuzzo, 2001). Stability of test scores over time reflect test-retest reliability, whereas split-half reliability assesses the internal consistency of items within the same measure. Interrater reliability of a measure can be determined through the assessment of amount of agreement between observers or judges (Fleiss, 1981; Shrout and Fleiss, 1979). As with validity, reliability is not a property of the instrument or test, rather reliability refers solely to the scores achieved from the administration of a measure, therefore, coefficients of reliability should always be reported when undertaking empirical research (Feldt & Brennan, 1989).

With regards to the present study, the assessment of the reliability of scores using independent observers was especially relevant. There are a number of reliability estimates used in determining interobserver reliability, including percentage agreement (for determining reliability based on dichotomous data), intraclass correlation coeffificents (ICC) which provide reliability estimates based on the use of two or more independent judges, and percent agreement estimates. The decision as to which is most applicable is dependant on the data and it has been suggested reliance on one method may not provide an accurate depiction of reliability, as each method can be problematic in certain situations (Rankin & Stokes, 1999).

Recent Psychometric Studies in Clinical Psychology

Three recent journal articles reporting on the psychometric properties of new measures and adapted measures reveal differences in the extent to which validity and reliability evidence is assessed and reported. These articles were found using the PsychInfo database and using the terms "psychometric evidence" and "new" and "measure". The three reviewed articles were included for review in the present study for two reasons. First, the articles presented results from validation studies of the measuresas, and second, the measures had a clinical focus.

McLeod and Weisz (2005), outlined measure characteristics of a newly developed scale used to assess child-therapist and parent-therapist alliance. Reliability evidence was presented in the form of interrater reliability (using ICC) and internal consistency (subscale interdependence). Face and content validity evidence was presented through outlining the development and refinement of the new scale, while convergent validity evidence was evidenced through correlation with an existing self-report alliance measure. Based on the review of reporting practices described earlier, these authors provided far more psychometric evidence than found in most articles reviewed.

In a second study, McIntosh, Jordan, McKenzie, Luty, Carter, Carter, et al. (2005) measured therapist adherence in psychotherapy for anorexia nervosa, using a modified version of the Collaborative Study Psychotherapy Rating Scale (CSPRS) with the symptom focus changed from depression to anorexia nervosa. Further revisions were made to reflect the approach of nonspecific supportive clinical management approach (NSCM). The reporting of psychometric evidence included estimates of interrater reliability using intraclass correlation coefficients and

coefficients of variation. Internal consistencey estimates were based on the calculation of Cronbachs alpha for the subscales and item-total correlations. Validity evidence was presented in term of discriminant validity between the subscales. While the results presented in this article provides far more psychometric evidence than found in the reviews discussed earlier, it is important to note that no evidence of face validity was presented.

Clark, Antony, Beck, Swinson, and Steer (2005) presented the results of a validation study of a newly developed measure used to screen for both frequency and severity of obsessive and compulsive symptoms in obsessive-compulsive disorder. Information regarding item construction showed face and content validity evidence. Validity evidence for the revised scale was presented for criterion-related validity (group comparison), discriminant and convergent validity (with existing measures), and factorial validity (confirmatory factor analysis). Reliability analyses included internal consistency, test-retest stability. The journal article presented results from an initial pilot study as well as a validation study of the revised measure. In all, the psychometric evidence presented in this journal article was far more than found in the reviews.

Overall, the degree of psychometric evidence presented in these validation studies show a much higher degree of reporting than found in earlier reviewsI. However, there are still differences in the amount and type of psychometric evidence reported. It is likely that this reflects differences in methodology and sample size, with the Clark et al. (2005) study being the most comprehensive in terms of reporting, while also including results from an initial pilot study and a validation study of the measure.

Summary

In short, determining and reporting the psychometric properties associated with any assessment instrument used in psychological research is, perhaps, one of the most important steps in the construction of a new measure. Further, the process of measure validation is an ongoing process, and as such, evidence for both the validity and the reliability of a measure need to be assessed and reported, a stance strongly recommended throughout literature on the reporting of psychometric evidence. The push towards the identification of specific process variables and their role in therapy outcome demands the tailoring of measures already in use, as well as the development of new measures. Three recent articles reviewed in the present study suggest that researchers are beginning to address issues in the reporting of psychometric evidence highlighted by a number of reviews and recommendations of reporting practices. In planning the present study, the recommendations for reporting psychometric evidence helped guide the choice of comparison measures as well as the analyses used to provide specific psychometric evidence.

CHAPTER 6

The Present Study

Overview

The present study was part of the ongoing research associated with the Cognitive Behavior Therapy Homework Project. The present study sought to further assess the psychometric properties of two new measures, one designed to assess therapist competence and adherence in homework integration (HAACS) and the other developed to ascertain client homework completion (HRS-II). This places the present study as contributing to the fourth objective of the Cognitive Behavior Therapy Homework Project, as outlined in Chapter 1.

As the present study evaluated the psychometric properties of these two new measures when used by independent observers, the research is presented as two separate studies. Study One provides the psychometric evaluation of the HAACS, while Study Two gathers psychometric evidence for the HRS-II. The following chapter presents the aims and hypotheses for the present research.

Aims

The present study had five discrete aims. The first was to examine the behavioural and cognitive theories of learning underlying the rationale for the use of homework in cognitive therapy, and was covered in Chapter 2. The second was to examine motivational theories related to client homework compliance and to link these theories to recommended practice for therapist homework integration. Within the second aim, covered in Chapter 3, was an evaluation of how homework compliance and specific therapist behaviours interact. The third aim followed on from

the second aim, in that, the assessment of therapist competence and adherence in administering homework has both research and clinical applications. Thus the third aim, which was addressed in Chapter 4, was to explore the assessment of therapist competence and adherence in cognitive therapy, with particular interest in how therapist behaviours in administering homework are assessed. The fourth aim was to examine the reporting practices of psychometric evidence for new or modified measures. These standards of reporting, covered in Chapter 5, formed the framework for the gathering of psychometric evidence in the present study. The fifth aim was to provide psychometric evidence for two new measures pertaining to homework, which is the focus of the remaining chapters.

In addition to the primary aims of the present study, there are three secondary aims, which will also be explored in the remaining chapters. It was noted previously, that research of therapist factors associated with treatment integrity (i.e., competence and adherence), have not explicitly focused on therapist adherence *and* competence in homework integration. More specifically, research has not addressed therapist homework integration in terms of adherence and competence as separate, but related aspects of treatment integrity. Further, there is little information regarding the temporal course of therapist competence and adherence in administering homework. Similarly, while client homework completion has been linked to positive outcome, little research has specifically looked at the temporal pattern of client homework completion across the phases of therapy. Therefore, the first secondary aim was to provide preliminary data regarding therapist differences in integrating homework into practice. Further, the second secondary aim was to explore the temporal nature of therapist adherence and competence in the administration of homework. Finally, the

third secondary aim was to gather preliminary data regarding client homework completion over time.

Specific Hypotheses for Study One

In relation to these aims, a number of hypotheses have been proposed. The specific hypotheses for Study One are outlined below.

Score Reliability of HAACS

Perepletchikova and Kazdin (2005) suggest that using independent raters skilled in the treatment delivery may be especially useful in the assessment of therapist adherence and competence. Further, it is suggested that using direct assessment methods, such as videotaping, may yield results less influenced by bias and demand characteristics. When assessing behaviour, it is necessary to determine whether or not independent observers are able to recognise when targeted behaviour occurs (Kaplan & Saccuzzo, 2001). While the HAACS has been constructed based on recommendations for the integration of homework into practice, the first step towards the validation of the measure is to determine whether it is possible to identify the presence of these behaviours in videotaped sessions. Thus, it was hypothesised that independent observers would be able to discriminate between occurring and non-occurring behaviours as assessed by adherence items of HAACS (hypothesis 1).

Further, it was hypothesised that independent observers would be able to agree on the occurrence of adherence behaviours (hypothesis 2). Given that the present study used raters skilled in the delivery of CBT, it was hypothesised that raters would provide consistent assessments of the degree of competence with which therapists enact these behaviours (hypothesis 3).

Internal Consistency of HAACS

The development of any scale requires an assessment of internal consistency of the measure (Kaplan & Saccuzzo, 2001; Kline, 1998). As described in Chapter 4, the HAACS has been constructed on theoretical and empirical foundations. Towards this, it was hypothesised that the adherence items for each section of the HAACS would be more closely related to their own section than to the other sections of the HAACS (hypothesis 4).

Correlational Analyses

Given the low associations between adherence and competence (Miller & Binder, 2002), it was hypothesised that adherence, as measured by the HAACS would show only low associations with measures of competence in homework integration (hypothesis 5). Further, it was hypothesised that HAACS competence scores would show strong, positive associations to other measures of homework competence (hypothesis 6). There were no specific hypotheses made concerning associations between therapist competence in homework and overall therapist competence or session integrity measures.

Together, the global ratings of the review, assign and design sections are representative of the overarching construct of competence rather than the specific criterion variables within each section (i.e., the specific behaviours targeted by the items). Therefore, when summed together, the global ratings of the review, design and assign sections of the HAACS are representative of overall therapist competence in homework. It was hypothesised that the sum of the HAACS global ratings would be positively associated with measures of overall therapist homework competence (hypothesis 7).

Evaluation of Therapist Differences

The very nature of therapy is responsive (Svartberg, 1999). This study intended to confirm the proposition made by Elkin (1999) that therapist variability in adherence occurs but does not produce statistically significant differences. Thus, it was hypothesised that therapist variability in adherence would occur (hypothesis 8). It was also hypothesised that therapist variability in adherence to homework integration behaviours would not be statistically significant (hypothesis 9). Further, as the assessment of competence measured by the HAACS treats competence differently than other measures of therapist competence (competence is assessed based solely on exhibited behaviours) it was hypothesised that there would be no difference between therapists in competence (hypothesis 10).

The Temporal Nature of Adherence and Competence

The therapeutic focus within cognitive behaviour therapy is not static. In fact, there are specific accepted phases of therapy (A. T. Beck, et al., 1979; J. Beck, 1994). The early phase involves the socialisation into the CT model, whereas the middle phase of therapy is focused on addressing cognitive change and the late phase of treatment is geared towards relapse prevention. With this in mind, it was hypothesised that there would be temporal differences in adherence, depending on whether or not the session occurred in early, mid or late therapy phase (hypothesis 11). There were no direct hypotheses as to where differences would occur. Further, it was hypothesised that there would be no temporal effect for competence (hypothesis 12)

Specific Hypotheses for Study Two

In relation to the aims of the present study, a number of hypotheses have been proposed. The specific hypotheses for Study Two are outlined below.

Score Reliability of HRS-II

The HRS-II was initially developed as therapist- and client-rated measures. However, research has noted the existence of rating bias and demand characteristics when relying on therapist and client ratings (D. Milne, 1987; Perepletchikova & Kazdin, 2005). Study Two included a psychometric assessment of the HRS-II when rated by independent observers. It was hypothesised that independent raters would be able to reliably rate the HRS-II (hypothesis 13).

Correlational Analysis

In order to provide concurrent validity evidence, it was hypothesised that total HRS-II scores would be positively correlated to ACRS scores, a measure of client homework compliance presently used in research (hypothesis 14). Given that the ACRS measures quantity of homework completion, it was hypothesised that there would be a strong, positive association between ACRS scores and scores for item 2 (quantity) of the HRS-II (hypothesis 15). In relation to the other items of the HRS-II, it was hypothesised that the individual HRS-II item scores would be positively correlated to each other (hypothesis 16) and to the ACRS scores (hypothesis 17).

Internal Consistency of HRS-II

As a newly developed measure, an assessment of the internal consistency of the HRS-II was required. The HRS-II was constructed based on the identification of

variables associated with homework completion. It was hypothesised (hypothesis 18) that factor analysis would result in a simple, interpretable solution consistent with the cognitive and behavioural determinants of homework completion described by Kazantzis and L'Abate (2005). Initial factor analyses of the HRS-II have produced two factors associated with therapist ratings of client homework completion and three factors based on client ratings of homework completion. There were two factors shared by client and therapist ratings: homework *benefits and completion*; *client beliefs* about the process of assigning/designing homework. Client ratings produced a third factor called homework *cost and completion*. While one of the aims of the study was to compare the results of the factor analysis from the present study to those of Bjornholdt (2006), no specific hypothesis was proposed in relation to this aim.

Client Variability in Homework Completion

While it generally assumed that there is client variability in homework completion, there has been no research relating patterns in homework completion across the phases of therapy. Given that the therapeutic focus changes throughout CBT (J. S. Beck , 1994), it was hypothesised that there would be differences in the pattern of client homework completion across the phases of therapy (hypothesis 19). No specific hypotheses were posited as to the pattern of homework completion.

CHAPTER 7

Method Study One

Session Data

This study used archived session data recorded on digital video discs (DVDs) from a recently completed replication study (Dimidjian, Dobson, Kohlenberg, Gallop, et al., in press). of the National Institute of Mental Health Treatment for Depression Collaborative Research Project (see Elkin, Shea, Watkins, Imber, Sotsky, Collins, et al. 1989) and a component analysis of Cognitive Therapy (see Jacobson, Dobson, Truax, Addis, Koerner, Gollon, et al. 1996). The study was conducted at the University of Washington

The full study treated 241 individuals between the ages of 18 and 60 years who met the criteria for Major Depression. Participants were randomly assigned to one of four treatment groups: Behavioural Activation (BA), Cognitive Therapy (CT), antidepressant medication, (ADM) or pill placebo. Severity of symptoms formed a stratification variable during the random assignment procedure. Clients in the BA and CT treatment groups received a maximum of 24 sessions over a 16 week period. Six experienced therapists provided BA and CT. Results of the study indicated that, for more severely depressed patients, BA was as efficacious as ADM, and more efficacious than CT. However, the authors noted that the CT group had a high level of nonresponders in the high severity group and the performance of CT in this study was, in part, due to this. Among less severe patients, CT provided better overall rates of response and remission rates as assessed by scores on BDI and HRSD than BA, ADM and pill placebo. However, the results indicated that there were no significant differences between treatments on the outcome measures.

As part of the original study's treatment integrity procedures, sessions were videotaped, with adherence to the treatment protocols being rated after the study was completed. Therapists were monitored for competence throughout. A more detailed report of clients, therapists and treatment procedures is presented in Dimidjian, et al., (in press).

It should be noted that treatment condition information was not supplied until after the present research was completed. The sample for this research consisted of both BA and CT sessions. The decision not to differentiate sessions by treatment condition was made based upon BA and CT sharing a number of elements (e.g., the structure of sessions, emphasis on collaboration, use of homework, etc.) being a shared element (Dimidjian, et al., in press). Information regarding the BA protocol can be found in published treatment manuals (Jacobson, Martell, & Dimidjian, 2001; Martell, Addis, & Jacobson, 2001). The provision of CT was consistent with standard CT for depression as specified by A. T. Beck, et al. (1979) and J. S. Beck (1995). Competency ratings, using the CTS, were supplied by Dr. Keith Dobson for all sessions used in the present study

Approval to use the archived data (including prior ethical approval and client consent) was obtained through Dr. Keith Dobson, one of the researchers involved in the originating study. The University of Washington IRB approved the protocol for the original study. The original study was supported by National Institute of Mental Health Grant MH55502 (R01), first to Neil S. Jacobson and, after his death, to David L. Dunner.

Independent Raters

Two registered clinical psychologists rated 74 sessions recorded on DVD. Both had postgraduate degrees in clinical psychology as well as specialist training in Cognitive Behavioural Therapy. One rater was an experienced clinical supervisor. The average clinical practice time for the raters was eight years.

Training

Prior to initial training, each rater received an information pack containing a copy of the chapter "A Guiding Model for Practice" (Kazantzis, MacEwan, & Datillio, 2005). Also enclosed in this package were copies of the measures used in the study and a laminated version of Therapist Quick Reference (TQR: Kazantzis, Deane & Ronan, 2005). The TQR is a summary the key therapist behaviours in recommending homework assignments along with the corresponding item number from the HAACS.

Training was conducted for eight hours over the course of a single day, with one booster session after four weeks of data collection. Dr. Nikolaos Kazantzis and a research assistant designed the training material. This material had been developed for a previous research project consisting of two pilot studies during the initial development of the Homework Adherence and Competence Scale. Further information regarding this material can be found in Wedge (2005). Training was provided by one of the developers of the HAACS as well as this researcher.

A total of two sessions were rated. After the each session was rated, the ratings were reviewed with any discrepancies in ratings of more than 2 points away discussed. Interrater reliability for the HAACS competence ratings was ICC (2,2) = .67, 95% CI .40 - .83. As another measure of interrater reliability, percentage agreement was computed, using the criterion expanded to include ratings within one point of each

other, producing a percent agreement for the HAACS of 84%. These reliability estimates are based on original ratings. After discussion, percentage agreement rose to 92%. Generating ICC for ratings of the THACS training sessions was not possible due to the number of items associated with each measure. Reliability coefficients are extremely sensitive to the length of a measure (Nunally, & Bernstein, 1994; Streiner, 2003a, 2003b). Rater agreement was calculated using percentage agreement, indicating 100% agreement between raters.

Measures

Homework Adherence and Competence Scale (HAACS; Kazantzis, Wedge & Deane, 2005)

The HAACS is a newly developed measure of therapist adherence and competence in the integration of homework into practice.The HAACS is a 19-item rating scale utilising a two-tiered hierarchy of adherence and competence. This allows for the interrelationship that exists between adherence and competence to be assessed in greater depth. Adherence items are scored as 0 (*no*) or 1 (*yes*). Item four, (*was an individualised conceptualisation used to make sense of any portion of non-completed homework?*) also has the potential to be scored as not-applicable (*n/a*).

Competence ratings are based upon the identification of items adhered to within each rated session. The anchors for the competence ratings vary for each item, but the underlying continuum for all items ranged from 1 (*poor*) to 6 (*excellent*). Each competence item has specifically developed anchors representing qualitative and quantitative difference between each point.

The HAACS consists of three sections based on specific therapist behaviours associated with homework review, design or assignment. This enables both adherence and competence to be assessed in association with the overarching aim of each section. At the end of each section is a global competence rating, which is scored based upon a general overall impression of the rated session. Global ratings are based on a seven-point likert scale, ranging from 0 (*non-adherence/extremely poor*) to 6 (*excellent*). For these ratings any special considerations, such as an increased emphasis on in-session practice for a client with a dependant interpersonal style, can be taken into account. Any factors found to influence the overall ratings can be recorded. It is possible then, that global ratings may differ from the competence ratings for the section.

Two previous pilot studies during the development of the HAACS found that post-graduate-level, clinical students were able to understand the behavioural anchors, as well as achieve a high level of overall rater agreement (79%; Wedge, 2005). Table 1 shows the adherence items for the scale.

The scoring for the two rating scales of the HAACS (adherence and competence) produces three index scores and three global ratings. As item 4 (*use of a situational conceptualisation for any portion of homework not completed*) has the possibility of a not-applicable adherence rating, the scoring procedure needs to reflect this. As some patients *do* complete all assigned homework, that particular therapist behaviour would not be required in that instance.

Table 1

Adherence items of the Homework Adherence and Competence Scale

Item Number	Item Description
Review	
1	Did the therapist discuss the completion of previously assigned homework?
2	Did the therapist provide verbal reinforcement (i.e., praise) for any portion of the homework carried out?
3	Was a situational conceptualisation (e.g., thoughts, behaviours, emotions, physiology) used in reviewing previously assigned homework?
4	Was an individualised conceptualisation used to make sense of any non-completd homework (i.e., linked non-completion to the client's automatic thoughts, underlying assumptions and rules, or core beliefs?
5	Did the therapist attempt to problem solve practical obstacles to the homework?
Design	
6	Was any new or revised homework discussed?
7	Did the therapist use any aspect of guided imagery to identify the client's coping strategies and beliefs related to the homework?
8	Did the therapist integrate a disorder-specific cognitive model with the individualised conceptualisation in designing homework?
9	Were homework tasks selected for comletion before the next session?
10	Did the therapist present any rationale for the homework?
11	Did the therapist ask about the client's ability and perceived difficulty of the homework?
12	Was any attempt made to facilitate in-session practice?
13	Did the theraoist use guided imagery to begin experiential learning for the homework in-session?
14	Did the therapist use a situational coceptualisation to help identify the client's beliefs and triggers (i.e., emotional, behavioural, physiological) for carrying out the homework in specific situations?
Assign	
15	Was there any attempt to summarise the rationale for the homework in relation to therapy goals?
16	Was there any attempt to specify how the homework will be practically integrated into the client's life (i.e., specification of when, where, how often, how long)?
17	Was there any consideration of potential difficulties for completinh the homework?
18	Was there any attempt to explain the outcome from the homework as having a learning 'experiment' focus?
19	Was there any attempt to summarise the homework?

Note. Item descriptions are based on the Therapist Quick Reference (TQR; Deane, Kazantzis & Ronan, 2005). The full HAACS measure can be located in Appendix A.

Adherence Index (AI). The HAACS is constructed so that therapist can receive ratings on a maximum of 19 items, each reflecting a critical aspect of the model for practice (see Kazantzis, MacEwen, & Dean, 2005). Behaviour is rated as 0 = *no* or 1 = *yes* based on whether or not the behaviour is observed. The adherence index is presented as a percentage score. That is, the total of observed behaviours divided by the maximum possible score of 19, multiplied by 100. However, when item 4 is rated as *not-applicable*, the maximum possible score for the scale is 18.

Session Competence Index (SC). Given the hierarchical relationship between adherence and competence constructs (for an overview, see Perepletchikova & Kazdin, 2005), the HAACS has been constructed to ensure that therapist ratings reflect competence on adherence behaviours. That is, those items rated "*yes*" on the adherence scale. Each therapist behaviour is rated in a scale ranging fro 1 through 6, with anchors for each rating clearly defined for the specific behaviour observed. Session competence is presented as a mean score based on the total HAACS ratings divided by the number of competence items rated.

Fidelity Index (FI). The hierarchical relationship between adherence, competence and fidelity constructs is reflected in the scoring procedure for fidelity index. The rationale for producing a fidelity index was account for both adherence and competence as well as not penalising therapists for not having been able to show competence using a situational conceptualisation in the review of homework (item 4). The fidelity index is the number of adherence items rated (plus 1 if Item 4 is rated as "*not applicable*") multiplied by the session competence index.

Therapist Homework Assignment Competency Scale (THACS; Bryant, Simons, & Thase, 1999)

The THACS is a four-item scale designed to assess therapist behaviours in administering homework. Each item is scored from 0 (very poor or not done) to 6 (very well done). The items domains include reviewing previous session homework, providing a rationale for homework tasks, assigning new homework and eliciting reactions and possible difficulties that might arise in doing new homework. The maximum score for this measure is 16. Based on the CTS, the THACS has provided limited data showing therapist competence in administering homework is a predictor of treatment outcome (Bryant et al., 1999). A copy of the THACS is included in Appendix B.

Cognitive Therapy Scale (CTS; Young & Beck, 1980)

The CTS is a measure of competence in cognitive therapy. The scale is used widely to provide supervisor ratings in training and certification programs (see Dobson, J. S. Beck, & A. T. Beck, 2005; Lau, Dubord, & Parikh, 2004; Sudak, J. S. Beck, & Wright, 2003; Weerasekera, et al., 2003) and has proven useful in providing a standardised method of rating therapy sessions in research (Sudak, J. S. Beck, & Wright, 2003). The CTS is an 11-item rating scale covering two domains; General Therapy Skills (e.g., agenda setting, interpersonal effectiveness, collaboration, pacing, use of feedback, and understanding); Conceptualisation, Strategy and Technique (e.g., focusing on key cognitions and behaviours, homework, strategy for change, and guided discover). Each item is rated using a 7-point Likert scale ranging from 0 to 6. The maximum score for the scale is 66, with 40 being the usual cut-score for defining competence CT in randomised clinical trials (Elkin, et al., 1999) and is the designated

passing score used by the Academy of Cognitive Therapy (Dobson, J. S. Beck, & A. T. Beck, 2005; Sudak, J. S. Beck, & Wright, 2003). Psychometric evaluations of the CTS have produced mixed results (for a review, see Kazantzis, 2003) and difficulty in determining Intraclass Correlation Coefficients (ICC: Vallis, Shaw and Dobson, 1986), although pooling data from two raters ihas been shown to increase the estimates of reliability. In clinical trials, sessions have been graded as satisfactory or unsatisfactory on the basis of CTS scores, with a score of 39 or less serving as the "red-line" (Elkin, 1999). The CTS is included in Appendix C.

Procedure

As this study had been designed, in part, to evaluate psychometric properties of the HAACS, sessions included for analysis in the present study were determined by matching completed CTS forms supplied by Dr. Keith Dobson, with labelled client digital video discs (DVDs). In this sample, 91 sessions were indicated as appearing on the DVDs. Of these, 12 did not appear on the labelled DVDs (13 %). Four sessions (4%) were not included for analysis due to technical faults occurring on the DVDs. One session (1%) was not included due to crisis intervention within that session. This left a total of 74 sessions for analysis.

Each rater received sealed envelopes containing collated rating forms (HAACS, THAACS, HRS-II and ACRS) and DVDs for each client. Raters viewed each client's selected sessions starting with the earliest session and progressing through the course of therapy. Ratings were made at the end of each session with the order of the scales to be rated to follow the order of collation. The raters viewed and rated the sessions separately in order to maintain independence of ratings.

Data Preparation

Therapist and client numbers from the original study were changed to ensure confidentiality in any reporting. Raw data for the HAACS measure was initially entered with as 0 (*no*), 1 (*not applicable*) and 2 (*yes*). This was necessary to determine the maximum score for the creation of the adherence and fidelity indices. Adherence data for yes and not applicable ratings were then recoded with not applicable as 0 and yes ratings as 1. There were no reverse coded items for any measure used in Study 1.

Prior to manipulation and statistical analyses, the data sets were checked for missing values. There were no missing values with the exception of THACS ratings, with three sessions not rated due to rater error. Although the missing values corresponded to zero ratings for the HAACS, it could not be assumed that the THACS ratings were, in fact, rated as zero. These particular sessions corresponded to rater disagreement as to the occurence of homework within the session, so it was decided not to impute a value for the judgement, as the raters were both qualified to determine whether or not homework was reviewed.

In order to carry out analyses, specific variable of interest were constructed. Therapist adherence to homework behaviours is operationalised as the HAACS Adherence Index, (HAACS AI). Therapist competence in homework behaviours is operationalised as the HAACS Session Competence Index (HAACS SC), the CTS homework item (CTS-HW) and the sum of the HAACS Global Ratings (GS). The combined construct of therapist adherence and competence in homework is operationalised by fullscale THAACS scores (THAACS). Session Fidelity Index (HAACS SF) provides the measure of fidelity to the overall model of homework integration. Overall therapist competence in cognitive therapy, was determined by full-scale CTS sum (CTS).

To assist with the analysis of the data as a function of the phase of therapy, sessions were assigned a code indicating whether or not the session occurred in the early, middle or late phase of therapy. In this sample, all rated sessions occurring after session 20 were identified by the raters as being penultimate sessions and were excluded from analyses incorporating phase of therapy. First sessions were also excluded from temporal analysis, as there would not be homework discussed. There was one session 2 excluded from temporal analysis, as the client had missed the initial session (this session was then treated as a first session). The rationale for excluding first and last sessions was that first sessions do no review homework, and final sessions generally do not design or assign new homework. Further, the exclusion of initial sessions is consistent with recommendations made by Tang and DeRubeis (1999); first sessions are very different from all other sessions, so they should be excluded from analyses. These exclusions resulted in sessions 2 through 20 being allotted codes for phase of therapy. Determining how to split the sessions into phases was difficult, as there is no set criterion in the literature for determining the exact sessions belonging to which phase.

Statistical Analyses Procedure

The data were analysed using the Statistical Package for the Social Sciences version 13 (SPSS; SPSS Inc., 2004). The data were assessed with the traditional tests of significance with $p = .05$, except where noted. In a recent review of therapist effects in the NIMH TDCRP and other psychotherapy studies, Crits-Critsoph and Gallop (2006) suggest that therapist effects from large, multisite clinical trials typically find small effect sizes for therapist effects, when using Cohen's (1988) definitions of effect size. Basing calculations for effect size in a naturalistic setting on results published by

Okiishi et al. (2003), Crits-Critsoph and Gallop suggest that the magnitude of therapist effect will likely be between small and medium. In the present study, the sample size to achieve 80% power to detect medium effect sizes would have required a sample upwards of 130. For this reason, any effect size for between group comparison that was equal to, or greater than Cohen's (1988) definition of a small effect was considered to be potentially important. Cohen offered the following conventions for effect size with the understanding that their interpretation should be made within the context within which it was entrenched: small: $d = .20$, mediun: $d = .50$, large: $d = .80$. For each F test, estimates of effect size were determined from sample means, or η^2 (eta squared) on the condition that $F > 1$ (Hair. Anderson, Tatham, & Black, 1998)

Score Reliability

In order for there to be confidence in analyses stemming from ratings made by the use of independent raters, there needs to be a high level of agreement between the raters. As a general rule, agreement between 60 to 74 percent is considered good, and agreement over 75 percent is deemed excellent (.Cichetti, 1994; Cichetti & Sparrow; Fleiss, 1981).

Reliability of adherence ratings was determined through a number of methods. Cohen's Kappa (κ) does not give an accurate depiction of agreement when observed frequencies of behaviour are low (Cichetti & Feinstein, 1990). Therefore, at an item level, reliability of adherence scores were supplemented with reliability estimates using overall percentage agreement from cross-tabulations. The resulting

data from the cross-tabulations were then used to calculate overall rater agreement indicating the percentage of times the observed ratings agreed.

Intraclass Correlation Coeffieicnts (ICC; Haggard, 1958; Shrout & Fleiss, 1979) provide an estimate of rater agreement that accounts for chance and systematic differences between raters (Shrout & Fleiss, 1979). While there are a number of models of ICC, Model 2 has been used for theses calculations and is available using the ICC function of SPSS version 13 (SPSS, Chicago, IL) Model 2 is generally chosen to show that a measure has broad applications (Shrout & Fleiss, 1979). Analysis was made by using a two-way design, which includes variance of both the scores and the raters in the calculations. This study used averages of the ratings, so the average measures option was chosen. Any systematic variation in scoring between the two raters was important to this study, so to account for this the absolute agreement subtype of ICC was used. The Shrout and Fleiss (1979) ICC(2,*k*) indicates a two-way, average measure calculation, where *k* is the number or raters, therefore the model used for ICC calculations in this study was ICC (2,2).

Tests for normality

The statistical assumptions of the tests used in the present study were checked prior to conducting the main analyses. Where appropriate, SPSS for Windows was used to validate these assumptions. To ensure that the data in the sample met with the assumption of normal distribution inherent in the planned analyses, mean scores and standard deviations were generated for the variables of interest. Univariate outliers were detected by first transforming the variables of interest to *z*-scores (Hair, Anderson, Tatham, & Black, 1998). For total observations less than 80, it is recommended that scores of 2.5 or greater should be investigated as outliers (Hair,

Anderson, Tatham, & Black, 1998). Upon review, it was decided that there were three extreme values and two other outliers, which were distinctive and were eliminated from formal analyses.

The Kolmogrov-Smirnov test (Lilliefors Significance Correction) is a statistical procedure used to assess whether or not data meets the criteria of normal distribution; the Kolmogorov-Smirnov test compares the observed results and compares them to what would be expected in a normal population (Hair, Anderson, Tatham, & Black, 1998,). When results of a Kolmogorov-Smirnov test is significant ($p < .05$), then the distributions are deemed different from a normal distribution. Where the significance value is $p > .05$, the distribution is not found to differ from a normal distribution.

Multivariate outliers were evaluated using Mahalanobis' distance (Hair, Anderson, Tatham, & Black, 1998). Two significant outliers were found in the present study and removed from multivariate analysis. Multivariate homogeneity of variance/covariance was assumed by Box's M test, which is extremely sensitive to the presence of non-normal variables and is considered significant $p < .001$ (Hair, Anderson, Tatham, & Black, 1998). Box's M is especially useful when the sample size is small and cell numbers unequal (Dancey & Reidy, 2002).

Internal Consistency

Due to the exploratory nature of the research and the small sample size in relation to the number of items, factor analysis of the HAACS was not possible as a method in determining internal consistency of the measure. Due to this, item analysis was undertaken in order to assess whether or not the separate scales were assessing separate parts of the homework process. That is, whether the items correlated best with their section more than the other two section. Nunnaly and Bernstein (1994).

consider item analysis as a precursor to factor analysis, and provides an initial indication of internal consistency of a newly constructed test.Item-total correlations were computed as a measure of internal consistency. Pearson product moment correlations were calculated between each HAACS item and the sum of the remaining items. Initially, each item was correlated with its own scale (with the item removed) and with the other homework scales (Green & Salkind, 2005).

Correlational Analysis

As part of the validation of the dimensions in therapist adherence, competence and fidelity in administering homework, the AI, SC, SI and GS scores were correlated with other measures of therapist competence: 1) competence as assessed by the Cognitive Therapy Scale (CTS), 2) homework competence as assessed by item 4 of CTS (CTS-HW), and 3) ratings on the Therapist Homework Assignment Competency Scale (THACS).Correlations were determined with the Pearson product moment correlation coefficient, *r*. The interpretation of the resulting correlation matrix used Cohen's (1988) conventions to determine the size of the linear relationship, with small $r = .10$, medium $r = .30$, and large $r = .50$.

Evaluation of Therapist Differences

To illustrate central tendency and variability information of therapists' competence in administering homework, means and standard deviations of HAACS session competence and adherence index for each therapist were computed. The overall therapist effects of the between-groups and within-groups factors were calculated and analysed using a one-way multivariate analysis of variance (MANOVA). The choice to use MANOVA over multiple ANOVAs was threefold.

First, MANOVA controls for Type I errors (Streiner, 1993), whereas the use of multiple ANOVAs increase the likelihood of a Type I error (Hair, Anderson,Tatham, & Black, 1998 ; Kline, 1998; Streiner, 1993). Secondly, the MANOVA procedure allows relationships between variables, that may not be apparent when analysed individually, become apparent (Streiner, 1993). Thirdly, MANOVA are not as sensitive to violations of the assumptions of homogeneity of variance and sphericity, which are problematic with ANOVA procedures (Hair, Anderson, Tatham, & Black, 1998; Streiner, 1993). MANOVA was conducted with Pillai's trace as the criterion rather than Wilks' Λ, as this is considered a more robust criterion when there are unequal cells in the analysis (Hair, Anderson, Tatham & Black, 1998). Due to variation in therapist representation in this sample, it was appropriate to use Pillai's trace as the criterion. Although it is recommended that cell size be a minimum of 20 to ensure normal distribution, it is possible to perform MANOVA procedures when cell sample sizes are smaller, so long as the sample in each cell is greater than the number of dependant variables included in the analysis (Hair,Anderson, Tatham & Black, 1998).

One-way analyses of variance (ANOVAs) were conducted on each dependant variable as follow-up tests to the MANOVA, using the Bonferoni method to control for Type I errors (Green & Salkind, 2005; Hair, Anderson, Tatham, & Black, 1998; Norusis, 2006). Post hoc analyses to the univariate ANOVA for adherence consisted of conducting pairwise comparisons.Each pairwise comparison was tested using the Bonferroni correction (Green & Salkind, 2005; Norusis, 2006).

The Temporal Nature of Adherence and Competence

Prior to undertaking analyses examining adherence and competence across the three phases of therapy, first and penultimate sessions were excluded from analysis. The remaining 56 sessions were then assigned to either early middle or late sessions. There is no definitive delineation of where these phases begin (see J. S. Beck, 1995; Feeley, DeRubeis, & Gelfand, 1999; Gilboa-Schechtman, & Shahar, 2006; Svartberg, 1999; Tang, Beberman, DeRubeis, & Pham, 2005), so the remaining sessions (sessions 2 though 20) were divided into thirds. That is, Early Phase (sessions 2 through 7), Middle Phase (sessions 8 through 13) and Late Phase (sessions 14 through 20). Means and standard deviations were calculated for each adherence and competence for each section of the HAACS for each phase of therapy. That is, Review adherence and competence, Design adherence and competence, and Assign adherence and competence.

The effect of the phase of therapy on therapist adherence and competence were analysed using a one-way MANOVA. The MANOVA procedure was described above. As follow-up, ANOVA procedure and post-hoc analyses were computed following the procedure in the assessment of therapist differences.

CHAPTER 8

Study One Results

Overview

This chapter reports the results of the present study in relation to the fifth aim, the psychometric evaluation of the HAACS when rated by independent observers. As part of the psychometric evaluation, a number of specific hypotheses were posited. The first hypothesis was that independent raters would be able to discriminate between occurring and non-occurring therapist behaviours, as assessed by the HAACS adherence items. The second, and related hypothesis proposed that there would be agreement by the raters on the occurrence of adherence behaviours. Similarly, it had been hypothesised that there would be agreement in competence ratings between the raters. It was further hypothesised that, as an indication of internal consitency of the HAACS, individual adherence items would be more closely associated with their own section rather than the other two sections of the HAACS. There were a number of specific hypotheses made about the expected associations between homework adherence and competence. Further, it had been hypothesised that there would be therapist differences in homework adherence but not in relation to competence. The final hypothesis was that there would be temporal differences in therapist adherence across the three phases of therapy.

Session characteristics

There were no missing data for the HAACS and CTS ratings. There were three unrated sessions due to rater error for THACS data (e.g., rater did not fill in ratings for THACS). The missing data were exluded listwise.

Table 2 outlines the frequency of therapist representation in the present sample. The sample consisted of 74 sessions meeting inclusion criteria (i.e., having a corresponding CTS rating), with 22 clients represented in these sessions. The mean number of sessions per client was 3.30 (SD = 1.36). There were six therapists in the sample, with one therapist representing almost half of the total sessions (47.3%).

Table 2

Frequency of therapist representation in the sample

Therapist	Frequency (n)	Percent
Therapist A	14	18.90
Therapist B	8	10.80
Therapist C	5	6.80
Therapist D	35	47.30
Therapist E	7	9.50
Therapist F	5	6.80
Total	74	100.00

Score Reliability of HAACS

In order for there to be confidence in analyses stemming from ratings made by the use of independent raters, there needs to be a high level of agreement between raters (Kaplan & Saccuzzo, 2001). Further, ratings based on observational methods assume that observers will be able to discriminate between occurring and non-occurring behaviours (Suen, 1988). As a general rule, rater agreement between 60 to 74 percent is considered good, and agreement over 75 percent is deemed excellent (Cichetti, 1994; Cichetti & Sparrow, 1981; Fleiss, 1981). As an initial check of rater

agreement, cross-tabulations were constructed for ratings between the two raters. This produced a total of 2812 pairings (i.e., 2 raters x 19 items x 74 sessions). Table 3 presents the cross-tabulations of the adherence ratings for the HAACS. Table 4 indicates the percent of between-rater agreement for rating responses.

Table 3

Rater by Rater Cross-Tabulation for HAACS Adherence Items

		Rater 1			
		Yes	N/A	No	Total
Rater 2	Yes	384	1	111	496
	N/A	1	20	17	38
	No	126	6	740	891
	Total	511	27	868	1406

Note. n = 74.N/A = Not Applicable ratings. HAACS= Homework Adherence and Competence Scale.

Table 4

Percent Between - Rater Agreement for HAACS Adherence Ratings

Rating	Between-Rater Agreement (%)[a]
Yes	27.30
Not Applicable	1.40
No	52.60
Overall	81.30

Note. n = 74. HAACS = Homework Adherence and Competence Scale.

[a] Between-rater agreement is based on percentage of total ratings, with overall percentage equal to the sum of percent agreement for yes, not applicable and no ratings.

The cross-tabulations in Table 4 show an overall percentage agreement of 81%, with Kappa = 0.62, $p < .001$. By convention, Kappa > .70 is considered acceptable interrater agreement (Fleiss, 1981). However, an important consideration is that Kappa is often confounded by low base rates of behaviour (Cichetti & Feinstein, 1990; Feinstein & Cichetti, 1990; Uebersax, 1987). Shown in Table 3, the frequency of agreed *yes* responses (n = 384) is considerably lower than agreed *no* responses (n = 740). This indicates the presence of an overall low base rate of behaviour, thereby raising concern about the use of Kappa as the sole indicator of interrater agreement in this instance. Due to the low base rate of behaviour indicated in the above cross-tabulations, reliability of adherence scoring at an item level was determined through overall percentage agreement from cross-tabulations. Intraclass Correlation Coefiecients (ICC) was used to show interrater agreement at an item level for competence ratings. Finally, overall rater agreement for the full scales of adherence and competence will be calculated using ICC.

Reliability of the Adherence Scores for the HAACS

To explore the reliability of the rater observations, cross-tabulation data were calculated on an item-by-item basis using raw data. Raw data were used to assess rater agreement on *not applicable* responses. The number of pairings for each item was 148 (i.e., 2 rater pairings x 19 items). The occurrence of adherence behaviours was indicated by absolute agreement on *yes* ratings (yes-yes). The percentage frequency for the behaviours indicates the frequency as a percentage of all observations. The resulting data from the adherence cross-tabulations were then used to calculate overall rater agreement indicating the percentage of times the observed ratings agreed (Yes-Yes, No-No, Not-applicable-Not-applicable). Together, frequency of behaviour and

percent agreement indicate the ability of raters to discriminate between occurring and non-occurring behaviours. Percentage agreement alone represents the degree of rater agreement on the occurrence of adherence behaviours Tables 5, 6 and 7 present the overall frequency of observed behaviours within and the overall rater agreement between the observations for that behaviour.

Table 5

Frequencies of Adherence Behaviours and Rater Agreement for HAACS Review Section

	Overall		
	Frequency of Behaviour		Rater
Therapist Behaviour (HAACS item number)	(n)	(%)	Agreement
Discuss the completion of previously assigned homework (1)	58	(78%)	92%
Provide verbal reinforcement for any portion of the homework carried out (2)	43	(58%)	86%
Use a situational conceptualisation to review previously assigned homework (3)	36	(49%)	82%
Use an individualised conceptualisation used to make sense of any portion of non-completed homework (4)	2	(3%)	**64%**
Problem solve practical obstacles to the homework (5)	3	(4%)	85%

Note: n = 74. Item descriptions are abbreviations of HAACS items taken from the Therapist Quick Reference (TQR; Kazantzis, Deane, & Ronan, 2005). HAACS = Homework Adherence and Competence Scale. Items shown in boldface indicate percentage agreement < 75%.

As shown in Table 5, rater agreement for adherence behaviours for the review section of was greater than 80 percent for most behaviours. Item 4, shown in bold, was rated with 64% overall agreement. However, this item had a "not-applicable" rating option, as it *is* possible for clients to complete all homework, making the use of an individualised conceptualisation for non-completion redundant. When *not-applicable* ratings were recoded as non-occurring behaviours (i.e., *no*), the rater agreement rose to 96 percent. This indicates that rater agreement was excellent in discriminating between the use and non-use of a situational conceptualisation for homework non-completion. It does, however, also show that in some cases, raters did not agree whether this behaviour was clinically appropriate (i.e., an individualised conceptualisation could have been discussed in the general course of therapy or skill review). The overall frequency of review behaviours ranges from 4% of all observations, through to 78% of all observations. These results suggest that raters were able to identify when review adherence behaviours occurred. Further, there was a high degree of concordance between raters when therapis adherence behaviours occurred.

Table 6 shows that seven of the nine HAACS items rating therapist behaviour in homework deign, achieved an excellent level of agreement between the raters (i.e., > 75%). However, using the conventions for evaluating levels of agreement (Cichetti, 1994; Cichetti & Sparrow, 1981; Fleiss, 1981) items 7 and 8 only achieved a good level of agreement, indicating some difficulty in discriminating when these behaviours occurred and did not occur. While two behaviours had high base rates of occurrence (i.e., > 65%), the overall frequency of the other eight behaviours was less that 35 percent, with item 13 (*use guided imagery to begin experiential learning in-session*) not observed by either of the raters.

Table 6

Frequency of Adherence Behaviours and Rater Agreement for HAACS Design Section

	Overall		
	Frequency of behaviour		Rater
Behaviour (HAACS item number)	(n)	%)	Agreement
Discuss new or revised homework (6)	54	(73%)	84%
Use guided discovery to identify coping strategies and beliefs (7)	16	(22%)	**73%**
Integrate a disorder-specific cognitive model with the individualised conceptualisation (8)	11	(15%)	**73%**
Collaboratively select task (9)	50	(68%)	79%
Present a rationale that aligns with client's treatment goals (10)	28	(38%)	84%
Ask about the clients' ability and perceived task difficulty (11)	9	(12%)	80%
In-session homework practice? (12)	9	(12%)	84%
Use guided imagery to begin experiential learning in-session (13)	0	(0%)	100%
Use a situational conceptualisation to identify beliefs and situational triggers (14)	1	(1%)	84%

Note. $n = 74$. Item descriptions are abbreviations of HAACS items taken from the Therapist Quick Reference (TQR; Kazantzis, Deane, & Ronan, 2005). HAACS = Homework Adherence and Competence Scale. Items shown in boldface represent percentage agreement < 75%.

The overall agreement for assigning behaviours, shown in Table 7, revealed that item 15 only achieved a good level of agreement between the two raters (i.e., 65%). The other behaviours achieved an excellent level of agreement, with percentage agreement over 80%. However, it should be noted that the overall frequency of assigning behaviours is quite low, with none achieving above 35 percent frequency. This suggests that although assigning behaviours were infrequent in this sample, the raters were able to discern these behaviours with a high degree of concordance.

Table 7
Frequency of Adherence Behaviours and Rater Agreement for HAACS Assign Section

	Overall		
Behaviour (HAACS item number)	Frequency	(%)	Rater Agreement
Summarise the rationale for the homework in relation to therapy goals (15)	6	(8%)	**65%**
Specify how the tasks will be practically possible (16)	19	(26%)	83%
Consider potential difficulties (17)	6	(8%)	81%
Emphasise learning 'experiment' focus (18)	9	(12%)	83%
Summarise the homework and obtain indication/ratings of readiness, importance, and confidence (19)	24	(32%)	87%

Note: n = 74. Item descriptions are abbreviations of HAACS items taken from the Therapist Quick Reference (TQR; Kazantzis, Deane, & Ronan, 2005). HAACS = Homework Adherence and Competence Scale. Items shown in boldface represent percentage agreement < 75%.

The overall reliability of the HAACS adherence scale was kappa= .77, indicating that the total HAACS adherence scale can be rated to a substantial level of agreement, based on guidelines for the interpretation of kappa suggested by Landis

and Koch (1977). As noted previously, the percentage agreement for adherence at the item level was excellent. There were items that were exceptions, achieving rater agreement in the fair to good range (Cichetti, 1994; Cichetti & Sparrow, 1981; Fleiss, 1981). As hypothesised, raters were able to discern when adherence behaviours occurred (hypothesis 1). Further, when behaviours occurred, raters agreed on the occurrance, as determined by the percentage agreement statistic (hypothesis 2).

In order to further estimate the reliability of the adherence ratings, Intraclass Correlation Coeffieicnts (ICC; Haggard, 1958; Shrout & Fleiss, 1979) were calculated. The ICC is a measure of agreement that accounts for chance and systematic differences between raters (Shrout & Fleiss, 1979). As described in the Statistical Procedures section in Chapter 7, there are a number of models of ICC, with ICC(2,*k*) indicates a two-way, average measure calculation, where *k* is the number or raters, therefore the model used for ICC calculations in this study was ICC (2,2).

ICC (2,2) calculations were computed for adherence data for the 74 rated sessions. When conducted on the raw data (not recoded), the overall ICC was .77, with 95% CI .74 to .79, indicating that there was only a 5% chance that the true degree of rater agreement did not fall between .74 and .79. When computed on recoded data (i.e., *not applicable* as *no*), the overall ICC remained the same, ICC= .77, with a 95% CI of .75 to .80, indicating that the not-applicable option did not effect overall rater agreement. Using guidelines for evaluating levels of ICC these results indicate that the total HAACS adherence scale can be rated at an excellent level of agreement (Cichetti, 1994; Cichetti & Sparrow, 1981; Shrout & Fleiss, 1981).

Reliability of the Competence Scores for the HAACS

As an indication of the reliability of competence scores at an item level and for the global ratings for each section, ICC (2,2) were calculated and are presented in Table 9. Item 13 (*use guided imagery to begin experiential learning in-session*) was not included in these analyses, as there were no competence ratings for the item due to the non-occurrence of this behaviour in this sample. The overall ICC (2,2) for competence ratings was .81, with a 95% confidence interval of .79 to .83. This indicates that the total HAACS can be assessed with an excellent degree of reliability (Cichetti, 1994; Fleiss, 1981) (hypothesis 3). As shown in Table 8, estimates of reliability at the item level were varied, and tended to be lower than the overall ICC estimates. Rater agreement for the global ratings ranged from .60 (95% CI = .23 - .77) to .91 (95% CI = .85 - .94) indicating that the global ratings demonstrated fair to excellent reliability. At the individual item level, reliability estimates lie along a continuum; seven items achieved an excellent degree of reliability (> .75), nine items fell within the fair to good range (.40 -.75), and three items achieved poor levels of reliability (< .40). As shown in bold in Table 8, there were a few negative values for ICC. It should be noted that ICC calculations are sensitive to variance in responses and can produce negative results when within-groups (intra-rater) variance exceeds the between-groups variance (McGraw & Wong, 1996). Therefore, the presence of low and negative ICC found in competence ratings do not necessarily suggest that these items are not reliable, but may indicate the presence of variability in within-rater responses for low base-rate behaviours. That is, reliability for ratings of low base-rate behaviours become problematic when even small variances occur (Thompson, 2003)

Table 8

Intraclass correlation coefficients for HAACS Competence Ratings

Item	ICC	95% CI	M	SD
1	.87	.79 - .92	3.09	1.81
2	.76	.58 - .86	2.00	1.60
3	.78	.65 - .86	2.00	1.76
4	.93	.90 - .96	.15	.75
5	.28	**-.12 - .54**	.26	.59
6	.71	.54 - .82	2.45	1.51
7	.59	.34 - .74	.85	1.08
8	.58	.32 - .74	.49	.73
9	.73	.57 - .83	1.66	1.06
10	.48	.18 - .67	1.15	1.02
11	.60	.37 - .75	.55	1.05
12	.80	.69 - .88	.37	.77
14	**-.13**	**-.80 - .29**	.18	.48
15	.35	**-.02 - .59**	.41	.64
16	.78	.65 - .86	.74	.98
17	.60	.37 - .75	.39	.83
18	.66	.46 - .75	.63	1.08
19	.79	.67 - .87	.47	.62
GR	.91	.85 - .94	2.35	1.63
GD	.65	.45 - .78	1.77	1.12
GA	.60	.23 - .77	.91	.92

Note. n = 74. HAACS = Homework Adherence and Competence Scale. GR = Global rating for review section of HAACS. GD = Global rating for design section of HAACS. GA= Global rating for assign section. Items shown in boldface indicate negative values for ICC.

Reliability of Ratings for Other Measures of Therapist Adherence and Competence

There were a number of competence measures included in this study to evaluate the validity of the HAACS. It was, therefore, an important consideration to determine the reliability of the scores on theses measures. No reliability analysis was possible for the CTS scores, as only one rater supplied the ratings. The CTS ratings came from a larger study (outlined in Chapter 7) and were provided by Dr. Keith Dobson for this research to provide comparison between therapist competence measures. Dr. Dobson is a recognised expert in scoring the CTS (Dimidjian, et al., in press). To determine the reliability for ratings on the THACS, intraclass correlation coefficient was generated, producing an ICC(2,2) =.72, with 95% CI of .58 to .81. Based on the general guidelines for reliability, this indicates a fair to good level of rater agreement (Cichetti, 1994; Cichetti & Sparrow, 1981; Shrout & Fleiss, 1981).

Tests for Normality

To ensure that the data in the sample met with the assumption of normal distribution inherent in the planned analyses, mean scores and standard deviations were generated for the various indices of the HAACS: Adherence Index (AI), Session Competence (SC) and Session Fidelity Index (FI). Mean scores and standard deviations were also computed for the CTS, CTS Homework item (CTS-HW) and THACS. These were examined for normality, resulting in five extreme scores excluded from analysis. This resulted in a total of 69 sessions included for bivariate and univariate analyses. Box and whisker plots revealed that the distributions were approximately normally distributed and there were no extreme scores. A visual examination of the distribution histograms indicated relatively normal distributions.

The Global Ratings for Review (GR), Design (GD), and Assign (GA) sections were also assessed for normality, and revealed a distinct variation from a normal distribution, with box-and-whisker plots revealing a strong positive skew. When the Global Ratings were summed together, the distribution became normal. Box and whisker plots revealed that these distributions were also normally distributed and that there were no extreme scores.

Kolmogrov-Smirnov tests were undertaken to further assure a normal distribution. No tests were significant ($p > .05$), indicating that all the distributions of the scores were not significantly different from a normal distribution.

As shown in Table 9, there was variability in therapist performance in administering homework, as can be seen in the range of scores. Also indicated in the table is that while HAACS SC achieved a higher maximum score than did CTS-HW, the mean rating of CTS-HW was higher than HAACS SC. This suggests that ratings for CTS-HW tended to be higher than those of HAACS SC. There was a large degree of therapist variability in adherence, indicated by the minimum and maximum ratings of AI which ranged from 5.26% to 69.44%. This variability in adherence is reflected in the HAACS FI, which achieved a mean rating of 18.33 out of a possible 114.

Table 9

Means, Standard Deviations, Minimum and Maximum Scores for Therapist Skill Ratings of Sample Sessions

Competence Rating	M	SD	Minimum	Maximum
HAACS AI[a]	35.86	16.45	5.26	69.44
HAACS SC[b]	2.58	0.70	1.17	5.17
HAACS FI[c]	18.33	8.91	2.50	37.85
HAACS GS[d]	4.91	2.60	0.00	10.00
Full-Scale CTS[e]	44.62	6.00	31.00	59.00
CTS-HW[f]	3.62	0.76	2.00	5.00
Full-scale THACS[g]	3.61	2.49	0.00	9.00

Note. n = 67, listwise deletion.HAACS AI= Homework Adherence and Competence Scale Adherence Index. HAACS SC = Homework Adherence and Competence Scale Session Competence. HAACS FI = Homework Adherence and Competence Scale Session Fidelity. HAACS GS = Homework Adherence and Competence Scale sum of global ratings. CTS = Cognitive Therapy Scale (fullscale). CYS-HW = Cognitive Therapy Scale Homework item. THACS = Therapist Homework Adherence and Competence Scale.

[a]Maximun score possible = 100
[b]Maximun score possible = 6
[c]Maximun score possible = 114
[d]Maximun score possible = 18
[d]Maximun score possible = 66
[f]Maximun score possible = 6
[e]Maximun score possible = 16

Internal Consistency of the HAACS using Item Analysis

Because the HAACS was a newly developed measure, it was desirable to gain some understanding of its' psychometric properties. To investigate the degree to which the HAACS three sections tapped specific aspects of homework integration, item-total correlations were computed as a measure of internal consistency. As a precursor to factor analysis (Kline, 1998; Nunnaly, & Bernstein, 1994), item analysis

provides an estimate of reliability based on the relationship of the individual item to the overall score (Norusis, 2006).

As the behaviour associated with item 13 was not observed in this sample, this item was not included in this analysis. Item analysis was conducted on the remaining 18 adherence items representing the recommended behaviours for homework integration. Transformation of scores was not necessary, as items were scored using the same metric. Using the HAACS scores from the 69 sessions included for analysis, Pearson product moment correlations were calculated between each HAACS item and the sum of the remaining items. Initially, each item was correlated with the other items in that section of the HAACS as well as being correlated to total adherence. The corrected item-total correlations are the Pearson correlations between the individual item score and the sum of the scores for the remaining items (Green & Salkind, 2005; Norusis, 2006); Nunnaly & Bernsein, 1994) The results of these analyses are shown in Tables 10.

In support of the measure's internal consistency, most items correlate more highly with their own section than with the other two sections. Results indicated that the presence of therapist review behaviours were not associated, to any degree, with the performance of design or assign behaviours. However, there was a moderate degree of correlation between therapist performance of designing and assigning behaviours, suggesting therapists performing design behaviours were likely to perform assigning behaviours. Similarly, the results suggest that therapists that performed assigning behaviours were also likely to have performed design behaviours.

Table 10

Item-total Correlations Between Individual HAACS Adherence Items and Each HAACS Section[1]

		HAACS Sections		
HAACS Items		Review	Design	Assign
Review Items				
1	Discuss the completion of previously assigned homework	**.83**	-.09	.12
2	Provide verbal reinforcement for any portion of the homework carried out?	**.77**	-.06	.17
3	Use a situational conceptualisation to review previously assigned homework	**.73**	.00	.13
4	Use an individualised conceptualisation used to make sense of any portion of non-completed homework	**-.03**	.04	.02
5	Problem-solve practical obstacles to the homework?	**.32**	.06	.08
Design Items				
6	Discuss new or revised homework	-.10	**.81**	.54
7	Use guided discovery to identify coping strategies and beliefs	-.01	**.46**	.52
8	Integrate a disorder-specific cognitive model with the individualised conceptualisation	-.20	**.47**	.25
9	Collaboratively select task	-.03	**.68**	.63
10	Present a rationale that aligns with client's treatment goals	-.19	**.68**	.54
11	Ask about the clients's ability and perceived tasl difficulty	.24	**.44**	.37
12	In-session homework practice?	.08	**.16**	.17
14	Use a situational conceptualisation to identify beliefs and situational triggers	.06	**.09**	.16
Assign Items				
15	Summarise the rationale for the homework in relation to therapy goals	.18	.56	**.44**
16	Specify how the tasks will be practically possible (i.e., when, where, how often, how long)?	.00	.47	**.52**
17	Consider potential difficulties	.22	.49	**.43**
18	Emphasise learning 'experiment' focus	.07	.41	**.30**
19	Summarise the homework and obtain indication/ratings of readiness, importance, and confidence	.13	.36	**.45**

Note. n = 69. HAACS = Homework Adherence and Competence Scale. Corrected item-total correlations shown in boldface.

[1] Although not a planned analysis, it was requested by Dr. Keith Dobson that multiple regression be performed in order to determine which adherence items best predicted overall adherence. This analysis is presented in Appendix F

Coefficient alphas were computed to obtain internal consistency estimates of reliability for the review, design and assign sections of HAACS (Review a= .70; Design a = .81; Assign a = .80. The overall alpha for the complete measure was a = .83. While the review section was relatively independent of the design and assign sections (r = -.03 and r = .17, respectively), there was a moderate degree of intercorrelation between the design and assign sections ($r = .67, p < .001$). However, the shared variance was 44.9%, leaving 55.1% variance unexplained. This suggests that while there is some association between the assign and design sections, there is still some degree of independence between the two sections. These results suggest that the separate sections show an acceptable level of internal consistency (hypothesis 4). However, there were three items with low item-total correlations, with $r < .30$, indicating that for this sample, these items were not associated with any degree to the other items in their corresponding section.

Correlational Analysis

As part of the psychometric evaluation of the HAACS, the HAACS competence and adherence indices were correlated with other measures of therapist adherence/competence: 1) competence as assessed by the CTS, 2) homework competence as assessed by CTS-HW, and 3) adherence/competence as measured by the THACS. These correlations provided an indication of the concurrent validity of the HAACS. The correlation matrix is presented in Table11. Adherence, as measured by HAACS AI was positively associated with measures of competence, but this association was weak for HAACS SC, CTS, and CTS-HW with explained variances of 3%, 4% and 18%, respectively. Correlations between HAACS AI and all other

variables produced large, positive, significant correlation coefficients. This indicates that adherence was not associated with competence to any degree (hypothesis 5).

Correlation between HAACS SC and THACS produced a significant, medium correlation, with 20% shared variance ($r = .45$, $I < .001$). Of interest is the low correlation between HAACS SC and CTS-HW ($r = -.05$), which indicates that there is virtually no relationship between these two scores (hypothesis 6). Correlations between THACS and CTS and CTS-HW scores indicated small to moderate associations. HAACS GS produced strong positive correlations with all variables except for CTS and CTS-HW, which produced small to moderate associations (hypothesis 7).

Correlations between THACS and all the other variables showed moderate to strong significant correlations with HAACS variables, and lower correlations with CTS variables. These results suggest that HAACS variables show a degree of concurrent validity with the THACS. On the contrary, the CTS showed low correlations between all measures except with the CTS-HW, which provided a strong positive association. This indicated that the HAACS and CTS measure competence differently, thus indicating that the HAACS

Table 11

Correlation Matrix for Measures of Therapist Adherence and Competence Variables[a]

Skill Rating		1	2	3	4	5	6	7
1.	HAACS AI	--						
2.	HAACS SC	.17	--					
3.	HAACS SF	.91**	.53**	--				
4.	HAACS GS	.85**	.53**	.94**	--			
5.	CTS	.21	.19	.28*	.24	--		
6.	CTS-HW	.42**	**-.05**	.35**	.35*	.58**	--	
7.	THACS	.83**	.45**	.92**	.93**	.22	.33*	--

Note. $n = 67$, listwise deletion. *$p < .05$,** $p < .001$. HAACS AI= Homework Adherence and Competence Scale Adherence Index. HAACS SC = Homework Adherence and Competence Scale Session Competence. HAACS SF = Homework Adherence and Competence Scale Session Fidelity. HAACS GS = Homework Adherence and Competence Scale sum of global ratings. CTS = Cognitive Therapy Scale (fullscale). CYS-HW = Cognitive Therapy Scale Homework item. THACS = Therapist Homework Adherence and Competence Scale.

Evaluation of Therapist Differences

One of the secondary aims for this study was to examine therapist variability in the administration of homework. More specifically, the examination of therapist variability in adherence to recommended homework-related behaviours, and the competence with which they performed these behaviours. In order to address this, it was decided to use the HAACS SC and HAACS AI as the dependant variables. The rationale for excluding other measures of therapist competence was that CTS scores reflect overall competence in cognitive therapy and not homework specifically.

Further, THACS scores combine adherence and competence in homework rather than separately. Given that the HAACS provides separate indices for adherence and competence, the use of the HAACS SC and HAACS AI provides a clearer indication of any difference. This is consistent with recommendations made by Perepletchikova and Kazdin (2005) in the determination of treatment integrity. The correlation between HAACS SC and HAACS AI (n = 67) was small and not significant, $r = .18$.

Means and standard deviations were computed for HAACS SC and HAACS AI for each therapist (see Table 12). The means and standard deviations for competence (HAACS SC) were relatively similar for all therapists. There was variation in therapist ratings of adherence (HAACS AI), with Therapists A, C and D receiving, on average, higher adherence scores than the other therapists. Therapist C showed competence ratings similar to the other therapists, but had the highest mean adherence rating. However, it should be noted that this therapist also had the smallest representation in the sample ($n = 4$).

Table 12

Mean Scores and Standard Deviations for Measures of Adherence and Competence as a Function of Therapist

		Measures			
		HAACS AI		HAACS SCI	
Group	Sessions (*n*)	M	SD	M	SD
Therapist A	(13)	42.21	17.39	2.68	.54
Therapist B	(7)	33.65	13.50	2.17	.28
Therapist C	(4)	49.01	2.17	2.86	.25
Therapist D	(31)	30.03	15.34	2.50	.56
Therapist E	(7)	46.16	14.28	2.33	.69
Therapist F	(5)	36.96	13.34	2.52	.49

Note. HAACS AI = Homework Adherence and Competence Scale Adherence Index. HAACS SC = Homework Adherence and Competence Scale Session Competence Index. Maximum possible score for HAACS AI = 100, HAACS SC = 6.

Prior to undertaking multivariate analysis of variance, the data were assessed for suitability for multivariate analysis. Multivariate outliers were evaluated using Mahalanobis' distance (Hair, Anderson, Tatham, & Black, 1998). Two significant outliers were found in the present study and removed from multivariate analysis. Multivariate homogeneity of variance/covariance was assumed by Box's M test, which is extremely sensitive to the presence of non-normal variables and is considered significant $p < .01$ (Hair, Anderson, Tatham, & Black, 1998). Independence of therapist behaviour was assumed as therapist adherence and competence in homework was not a focus of the originating study.

One-way multivariate analysis of variance (MANOVA) was used to evaluate the effect of individual therapist efforts on adherence and competence in homework. It had been hypothesised that there would be differences in therapist adherence (hypothesis 8) but that therapist competence would remain stable (hypothesis 9). Box's M test was not significant ($p > .01$), indicating that the assumption of equality of variance/covariance was not violated. The results indicated a large main effect for therapists on the dependant measures, HAACS SC and HAACS AI, Pillai's Trace = .31, $F(10, 122) = 2.2$, $p < .025$. $\eta^2 = .15$, $d = .84$. The multivariate η^2, or eta squared, provides an indication of observed power. In this instance, η^2 based on Pillai's trace was .15, with an observed power of .84.

One-way analyses of variance (ANOVA) were conducted on each dependant variable as follow-up tests to the MANOVA. Using Bonferroni correction to control for Type I errors, each ANOVA was tested at the .025 level. The results indicated a large main effect for therapist on adherence, $F(5, 61) = 3.34$, $p < .02$, $\eta^2 = .22$, $d = .80$ (hypothesis 8). Results showed no significant therapist differences for competence (hypothesis 9).

Post hoc analyses to the univariate ANOVA for adherence consisted of conducting pairwise comparisons to determine which therapist affected adherence most strongly. Each pairwise comparison was tested at the .025 divided by 8, or the .004 level. There were no significant differences between therapist found in these analyses (hypothesis 10). [2]

[2]It has been proposed by Crits-Critsoff and Mintz (1991) that therapist effects could be noted when using a p value of .2.or .3. With the initial $p = .2$, and using the Bonferroni method to protect against Type I errors, the result of the ANOVAs were unchanged, with therapist differences only associated with adherence. Post hoc analyses were tested at the .02 level (.1 divided by 6). No significant differences between therapists were found.

The Temporal Nature of Adherence and Competence

As the first step in undertaking analyses examining adherence and competence across the three phases of therapy, first and penultimate sessions were excluded from analysis. The remaining 56 sessions were then assigned to either early middle or late sessions. There is no definitive delineation of where these phases begin (see J. S. Beck, 1995; Feeley, DeRubeis, & Gelfand, 1999; Gilboa-Schechtman, & Shahar, 2006; Svartberg, 1999; Tang, Beberman, DeRubeis, & Pham, 2005), so the remaining sessions (sessions 2 though 20) were divided into thirds. That is, Early Phase (sessions 2 through 7), Middle Phase (sessions 8 through 13) and Late Phase (sessions 14 through 20). The frequency distribution of the specific sessions for each phase included for analysis is presented in Table 13.

Table 13

Frequency of Sessions from Three Phases of Therapy

Early Phase	(n)	Middle Phase	(n)	Late Phase	(n)
Session 2	7	Session 8	6	Session 14	2
Session 3	3	Session 9	2	Session 15	4
Session 4	1	Session 10	5	Session 16	2
Session 5	3	Session 11	3	Session 17	2
Session 6	2	Session 12	2	Session 18	3
Session 7	1	Session 13	3	Session 19	3
				Session 20	2
Total	17		21		18

Note. Early phase included sessions 2 through 7. Middle phase included sessions 8 through 13. Late phase included sessions 14 through 20. Total number of sessions, $n = 56$.

In order to examine differences in adherence and competence across phases of therapy, scores reflecting both the section of the HAACS (i.e., review, design and assign) and the phase (i.e., early, middle and late) were computed. The means and standard deviations for adherence and competence section ratings are shown in Table 14. The pattern for adherence ratings across therapy phases indicates that during the early phase, therapists exhibited more adherence behaviours than in later phases of therapy. Competence ratings were lower for assign behaviours across all phases of therapy, and these ratings showed the largest variation during the middle phase of therapy. On average, therapists demonstrated higher competence ratings on review behaviours across all phases of therapy.

Table 14

Means and Standard Deviations of HAACS Adherence and Competence Behaviours Over the Phases of Therapy

			Therapy Phase					
	Overall		Early		Middle		Late	
Section	M	SD	M	SD	M	SD	M	SD
HAACCS RA	2.44	1.09	2.85	.79	2.62	.91	1.86	1.33
HAACCS DA	3.42	1.68	4.18	1.44	3.24	1.72	2.92	1.68
HAACCS AA	1.41	1.22	1.82	1.29	1.19	1.25	1.27	1.07
HAACS RC	2.96	1.16	3.11	.73	3.30	.90	2.44	1.56
HAACS DC	2.20	.86	2.33	.76	2.21	.93	2.18	.92
HAACS AC	1.37	.96	1.57	.63	1.19	1.16	1.47	.97

Note. Early sessions, n =17 . Middle sessions, n = 21. Late sessions, n = 18. HAACS RA is adherence for review section of HAACS. HAACS DA is adherence for design section of HAACS. HAACS AA is adherence for assign section of HAACS. HAACS RC is competence for review section of HAACS. HAACS DC is competence for design section of HAACS. HAACS AC is competence for assign section of HAACS. HAACS = Homework Adherence and Competence Scale. Maximum possible scores for variables are: RA = 5, DA = 9. AA = 5. All Competence ratings = 6.

Table 15 displays the correlations between the variables. As shown, there were three significant, positive correlations. Review adherence and competence scores showed a moderate, positive association (r = .66), indicating that when therapists were more adherent to review behaviours, they tended to do so more competently. There was a significant weak association between ratings of design and assign adherence (r = .38). Similarly, design and assign competence scores were weakly and significantly related (r= .36). These results indicate that therapists exibited similar rates of adherence for both design and assign behaviours, and they tended to do so with similar degrees of competence.

Table 15

Corerelation Matrix of HAACS Review, Design and Assign Sections for Adherence and Competence

		1	2	3	4	5	6
1	HAACS RA	--					
2	HAACS DA	.05	--				
3	HAACS AA	.07	.38*	--			
4	HAACS RC	.66**	.01	.12	--		
5	HAACS DC	.12	.21	.19	.22	--	
6	HAACS AC	-.07	.30	.15	-.63	.35*	--

Note. n = 56 HAACS RA is adherence for review section of HAACS. HAACS DA is adherence for design section of HAACS. HAACS AA is adherence for assign section of HAACS. HAACS RC is competence for review section of HAACS. HAACS DC is competence for design section of HAACS. HAACS AC is competence for assign section of HAACS. HAACS = Homework Adherence and Competence Scale.

Prior to undertaking multivariate analysis of variance, the data were assessed for suitability for multivariate analysis using the same criteria described in the assessment of therapist variability. It had been hypothesised that there would be a differential effect on the three sections of the HAACS for adherence across the three phases of therapy (hypothesis 11). Further, it had also been hypothesised that there would be no phase of therapy effect for therapist competence (hypothesis 12).

To examine possible differences in therapist adherence and competence as a function of therapy phase, a one-way multivariate analysis of variance (MANOVA) was conducted with phase of therapy as the dependent variable. Box's M test was not significant ($p > .01$) indicating that the assumption of the equality of variance/covariance was not violated. The results indicated a large main effect for phase of therapy, Pillai's trace = .39, $F(12, 96) = 1.96$, $\eta^2 = .20$, $d = .89$. Univariate tests using Bonferroni correction revealed that there was a large main effect for review adherence, $F(2, 52) = 5.36$, $p = .006$, $\eta^2 = .18$, $d = .84$, but not for design or assign adherence. Post hoc analyses using Bonferroni correction revealed no significant difference for review adherence due to phase of therapy.

CHAPTER 9

Discussion Study One

Overview

This study set out to provide psychometric evidence of the HAACS, a newly developed measure of therapist adherence and competence in the administration of homework. Additionally, there was an investigation of therapist variability in homework adherence and competence as well as an exploration of the impact of therapy phase on therapist competence and adherence in homework integration. This chapter presents a discussion of the main findings. The limitations of the study and the implications for future research and clinical practice will be presented in Chapter 13, General Discussion and Conclusions.

Reliability

The HAACS was developed to assess therapist competence and adherence in the administration of homework in CBT. To be considered useful in clinical practice, a measurement instrument must have adequate psychometric properties, particularly in relation to rater reliability. Data from this study suggest that, overall; the HAACS is used with an excellent degree of reliability for both adherence (.77) and competence (.81) when rated by independent observers.

With regard to common measures of observer reliability, percentage agreement and kappa treat systematic observer bias as error, contingency coefficients and the product–moment correlation do not, and ICC calculations may or may not treat observer bias as error, depending on whether interrater reliability is assessed using a consistency or absolute agreement definition of correlation (McGraw & Wong,

1996; Suen, 1988).

Within the context of behavioral assessment, such as the assessment of therapist adherence and competence, reliability indexes provide a means by which to evaluate the effectiveness of observer training and the objectivity with which target behaviors are measured (Berk, 1979). As the present study relied upon independent observer ratings, it first had to be shown that the observers were able to identify whether targeted therapist behaviours occurred in the viewed sessions. Cross-tabulation between the two raters showed a high level of percentage agreement for the occurrence of adherence behavioiurs (81.3%), with an overall estimate of rater agreement, kappa = .62. Although this indicates an unacceptable degree of agreement according to Fleiss (1981), the low base rate of behaviours in the present study warrants caution when interpreting kappa estimates. It has been noted that kappa is affected by low base-rates, and in instances of low base rate, kappa values may not necessarily reflect low rates of overall agreement. The paradox of high percentage agreement coupled with low values of kappa has been detailed extensively (for an overview, see Cichetti & Feinstein, 1990; Feinstein & Cichetti, 1990). It is likely that, in the present study, the kappa paradox was present.

It had been hypothesized that raters would be able to discern between occurring and non-occuring behaviours (hypothesis 1). When assessing rater agreement at the item level, the low base-rate of some therapist adherence behaviours was apparent. Only four of the 19 adherence behaviours occurred in more than 50% of the sessions, with eight behaviours being observed in less than 20% of the sessions. One adherence behaviour, *use guided imagery to begin experiential learning in-session* was not observed by the raters in any session. These results indicated that either the therapists did not perform the behaviours and was correctly identified by the raters, or that these

behaviours did occur, but were not recognised as such. However, the raters tended to agree on whether behaviours were and were not observed in-session, suggesting that raters were able to reliably discriminate between occurring and non-occuring behaviours. Further, ratings were made by experienced clinical psychologists with accredited specialist training in CBT and had undergone shared training in the use of the measure, These factors addressed concerns voiced by Jacobson (1998) regarding the reliability of some ratings on competence measures, such as the CTS, where expert raters did not undergo shared training.

Due to the low base rates of some adherence behaviours, estimates of reliability were based on percentage agreement. It had been hypothesized that raters would be able reliably rate adherence behaviours (hypothesis 2). In support of this hypothesis, the data suggested that although some therapist behaviours occurred infrequently, raters tended to agree, overall, on ratings of adherence. With few exceptions, estimates of reliability based on percentage agreement were greater than the 75% minimum agreement recommended by Fleiss (1981).

It should be noted that there were a few concerns raised about the reliability of ratings for two adherence behaviours. The first related to ratings for item 4 (*use an individualised conceptualisation used to make sense of any portion of non-completed homework).* This is the only behaviour in the HAACS with a *not applicable* option for adherence ratings. The reason for this option was that it *is* possible that some clients may complete all assigned homework, thus non-completion would not need to be addressed. However, it was apparent in the data, that while raters tended to agree when the behaviour did not occur, there was disagreement between raters as to whether this was clinically appropriate. The second item with lower rater reliability was item 15 (*summarise the rationale for the homework in relation to therapy goals*). The rater agreement was only 64% for

this item, which suggested that it was possible that the raters had difficulty identifying the behaviour. Another possibility is that the therapists did summarise the homework, but the raters did not agree as to whether this was done in relation to the goals of therapy. It is possible that the reliability of the ratings for these two behaviours may be a training issue, which suggests that in future, special attention be given to these two items to ensure better agreement on ratings for these two behaviours.

It had been hypothesised that therapist competence in homework integration could be reliably rated by independent observers (hypothesis 3). Reliability estimates at the item level for competence, using ICC, ranged from -.13 to .93, with all but three items attaining estimates of reliability greater than .60. Estimates of reliability for the global ratings ranged from fair to excellent, with the design and assign estimates achieving only a fair degree of reliability. While the results appear varied, it should be noted that ICC estimates of reliability using the absolute agreement option tend to result in lower estimates of reliability (Hall, Groome, Streiner, & Rochon, in press; McGraw & Wong, 1996). Further, within-rater variability becomes problematic for ICC calculations when ratings of are made on low base-rate behaviours, as even small variances may impact greatly on the results (Thompson, 2003). However, the results for interrater reliability for competence are comparable to results reported for competency ratings using the CTACS (Barber, Liese & Abrams, 2003).

Overall, interrater reliability from the current study compares favorably with the results from widely used measures of adherence and competence (the CSPRS and the CTS, respectively; DeRubeis & Feeley,1990; Hill et al., 1992; Hollon et al., 1988; Vallis et al., 1986). Thus, the HAACS is able to be reliably rated on both adherence and competence by indepentent observers.

Internal consistency

While the HAACS is divided into three separate sections, each containing behaviours specific to the section (e.g. review, design and assign), data indicate that the design and assign section items are somewhat homogeneous. Item-total correlations indicated that items in the design and assign sections load moderately to high on both sections, with both scales sharing approximately 45% of their variance.

Given the existence of longstanding recommendations for designing and assigning homework (e.g., A. T. Beck et al., 1979), it is surprising that these two sections are not more differentiated in their assessment of adherence. It may be that the level of abstraction in the HAACS design and assign sections may serve to make the judgements on various items nonindependent (see Vallis, Shaw, & Dobson, 1986 for a discussion of similar issues with the CTS). Previous measures of therapist adherence and competence in homework integration have not included specific items relating to the actual design component of homework administration. It may be that while the designing and assigning of homework are conceptually and theoretically different, they may not be independent in practice. That is, when homework is designed in session, it is then assigned. It is possible that the assignment and designing of homework occurred at the same time in this sample, which resulted in the lack of independence between the design and assign section. The HAACS is not designed to assess the specific timing of homework behaviours so it is unclear from the data if this was the case in this sample.

The results do not suggest that the design and assign sections be combined, as this is the first study using the measure. What the results do suggest though, is that, for this sample, there was an increased likelihood of assigning homework when homework was designed in-session. It is possible that the use of archived data, where homework

was not the focus of investigation, may have influenced the findings somewhat. It may be, that when a specific homework protocol is used, there may be greater independence between the design and assign sections of the HAACS.

Correlational Analysis

It had been hypothesised that adherence, as measured by the HAACS would show only low associations with measures of competence in homework integration (hypothesis 5). Although adherence is a prerequisite for competence (Nezu & Nezu, 2004; Perepletchikova & Kazdin, 2005), conceptually, it should be possible to distinguish between the two. The results from this study indicate that, as hypothesised, the scoring of the HAACS is able to differentiate between adherence and competence. Correlations revealed only a small but not significant correlation between adherence and competence as defined by the scores of HAACS Session Competence Index and the HAACS Adherence Index. This is contrary to the assertion that adherence and competence are associated when rated by the same judges (Barber & Crits-Critsoff, 1996). Further, this differs from findings for adherence and competence correlations reported by Barber and Crits-Christoph (1996), Barber, Liese and Abrams (2003), and Shaw et al. (1999).

It had been hypothesised that HAACS competence scores would show strong, positive associations to other measures of homework competence (hypothesis 6), but the results did not support this hypothesis. HAACS session competence showed a moderate degree of association with THACS ratings, and virtually no association with competence in homework, as assessed by the homework item of the CTS (r = -.05).

These results suggest that the HAACS session competence measures therapist competence in homework integration differently than the homework item of the CTS.

It had further been hypothesised that the sum of the HAACS global ratings would be positively associated with measures of overall therapist homework competence (hypothesis 7). To this end, the results support this hypothesis. Overall homework competence, based on the summed global ratings of the HAACS showed a significant, but weak, association with CTS homework item but an almost perfect, significant association with THACS ratings. HAACS measures of adherence and session fidelity did show a moderate, positive correlations with CTS homework item, and significant strong positive associations to THACS ratings.

While the results indicated that the HAACS session competence ratings effectively isolated competence from adherence, this produced a paradox regarding the interpretation of the findings. The HAACS session competence measures therapist homework competence, and should therefore, have shown some degree of association with CTS homework item, especially when HAACS session competence was moderately correlated to the THACS, which was based on the CTS homework item. One possible explanation is that using a different rater for the CTS ratings of homework competence may have confounded results due to that rater using significantly different rating criteria. Another possibility for the results lies in the scoring procedure of the different measures. Specifically, the score for session competence of the HAACS is computed solely for items adhered to, and does not indicate the overall degree of therapist adherence to specific homework behaviours. The overall session competence is the mean competence exhibited by a therapist in the execution of homework integration behaviours. This represents a marked departure from the scoring procedures of other measures of therapist competence that

sum competence ratings. The rationale for this departure is that, by definition, competence is the degree of skill with which a therapist delivers therapy. The scoring of the HAACS treats competence as an overall skill level and not as an increasing construct. It is possible that the procedure of summing competence may explain the existence of high correlations between competence adherence found by Barber and Crits-Christoph (1996), Barber, Liese and Abrams (2003), and Shaw et al. (1999). That is, when competence is summed, this may result in a competence score reflecting both adherence and competence rather than competence alone.

Although the interpretation of the results of the correlational analysis provided mixed results in terms of the hypotheses, it appears that the HAACS shows concurrent validity with the THACS, another measure of therapist competence in homework integration. Further, suggested by the results is that the HAACS measures different aspects of therapist competence than assessed by the CTS.

Evaluation of Therapist Differences

It had been hypothesised that therapist variability in adherence would occur (hypothesis 8). It was also hypothesised that therapist variability in adherence to homework integration behaviours would not be statistically significant (hypothesis 9). Adherence in this study did show variability across therapists, which is consistent with the notion of therapy as being responsive rather than simply following an inflexible protocol.

It had further been hypothesised that there would be no difference between therapists in competence (hypothesis 10).The results indicated that, overall, there was little variation in therapist competence in this sample. This suggested that competence was fairly homogeneous across therapists in the present study. While this should be

expected as the data originated from a randomised clinical trial, this is contrary to findings elsewhere suggesting that therapist competence varies to such an extent as to make analysis impossible (for example, see Elkin, 1999).

It should be noted that the results concerning therapist variability are only preliminary, and there are a number of qualifiers that need to be addressed. First, the results only suggest that there is a pattern of difference in adherence based on therapist effect. The results did not rule out the possibility of differences between therapists due to which therapy was provided by the therapist, either behavioural activation or cognitive therapy. This may account for the finding of significant differences between therapists in adherence. Second, the specific effects of client characteristics on therapist adherence could not be ruled out in this study. It is possible that client characteristics such as symptom severity may have influenced the degree of adherence shown by the therapists.

The Temporal Nature of Adherence and Competence

It had been hypothesised that there would be temporal differences in adherence behaviours, depending on whether or not the session occurred in early, mid or late therapy phase (hypothesis 11). The results indicated that therapist designing and assigning behaviours did not differ significantly across the three phases of therapy. However, the results suggest that therapist review behaviours do vary in relation to the phase of therapy. However, analyses did not identify where this difference occurred. Comparison of the means in Table 14 suggest that the most likely difference occurs in relation to the early sessions versus late phase of therapy.

Although the results suggested the presence of a temporal change in the pattern of therapist review adherence behaviours, a number of confounds cannot be ruled out. In particular, there was an unequal representation of therapists in this sample, and this may account for the temporal differences found in these analyses. Further, the sample did not include sufficient client-therapist pairings across the phases of therapy to sufficiently control for client effects. While the results indicate an effect on review adherence due to the position of the session in the phase of therapy, the client effects could also not be ruled out. The lack of access to outcome measures precluded the inclusion of client factors such as symptom severity in this analysis, so it remains possible that the results may indicate the effect of changes in client functioning on therapist adherence behaviours rather than differences due solely to a temporal effect.

Further, it was hypothesised that there would be no temporal effect for competence (hypothesis 12). The findings of no significant variation in competence ratings across sessions may indicate a conclusion that the assessment of competence does not need to be ongoing. However, in small samples such as the present, it would be inappropriate to draw inferences concerning a nonsignificant effect (i. e., the variance components). It is important to emphasize that the need for a larger sample cannot be obviated, as this would yield more precise estimates. The finding of no significant differences in competence across therapy is in line with findings reported by Barber and Crits-Christoph (1996) and Hill, O'Grady, and Elkin (1992).

CHAPTER 10

Method Study Two

The overall method and procedures for Study Two are the same as for StudyOne, and are outlined in Chapter 7. Results from rater training in the use of the measures for Study Two are described in this chapter. The measures used for Study Two are described below along with the data analytic procedures relevant to this study.

Rater Training

Rater training in the use of the HRS-II was conducted at the same time as training in the use of the HAACS. Prior to viewing training sessions, all measures were reviewed item by item to ensure rater understanding of the concepts covered by each item.

Two sessions were rated during training, with an overall interrater reliability for the HRS-II of ICC (2,2) = .56, 95% CI -.09 - .79. Although this is less than the recommended .70 (Fleiss, 1981), a low ICC estimate does not necessarily indicate poor overall agreement, but may reflect the uses of the absolute agreement option, which is a more stringent criteria for estimating rater agreement (Hall, Groome, Streiner, & Rochon, in press). Prior to training, it was determined that scores within one point of each other would indicate rater agreement. Based on this criteria, rater agreement was deemed acceptable, as all HRS-II item ratings were within one point of each other. Differences in ratings were discussed, resulting in consensus ratings indicating 100% agreement. Rater agreement for the ACRS was 100%, with no differences in ratings made by the two raters.

Measures

Homework Rating Scale-II (HRS-II; Kazantzis, Deane & Ronan, 2005)

The HRS-II is a 12-item scale designed to measure the theoretically meaningful determinants of homework completion. There are two forms available, one for therapist ratings of the client and the other a client self-report measure. The therapist rated version will be used in this study and can be viewed as analogous to supervisor ratings. Early indications suggest the existence of three factors when rated by both therapist and client (Bjornholdt,2006). The factor structure has found to be similar among client ratings, although the presence of a third factor when used by clients is not so clearly defined. Table 16 outlines the results of Principal Component factor analysis using oblique rotation for therapist and client ratings of homework completion.

Table 16

Therapist and client factors for the HRS-II identified in previous research

Therapist Factor 1[c]	Therapist Factor 2[d]	Therapist Factor 3[b]	Client Factor 1[c]	Client Factor 2[e]	Client Factor 3[d]
Pleasure	Rationale	Difficulty	Pleasure	Difficulty	Rationale
Mastery	Comprehension		Mastery	Obstacles	Comprehension
Progress	Collaboration		Progress	Quantity[a]	Collaboration
Quantity	Specificity		Quantity	Quality[a]	Specificity
Quality	Goals		Quality		Goals
Obstacles[a]					

Note. These results are from Bjornholdt (2006).
[a] these items achieved factor loadings less that .50.
[b] Therapist Factor 3 does not have the minimum required three items.
[c] Described as homework benefit and completion.
[d] Described as client's beliefs about the process of assigning/designing homework.
[e] Described as homework cost and completion.

The HRS-II is a 12-item measure consisting of items found to be theoretically and empirically associated with homework completion (Kazantzis, Deane & Ronan, 2005). Initially developed as a client and therapist rated measure, the potential for its' use in research and clinical supervision has led to the inclusion of an observer rated option. The measure was designed to incorporate cognitive and behavioural theories of factors involved in homework completion (Kazantzis, Deane, & Ronan, 2005; Kazantzis, Ronan, & Deane, 2004). Items are rated using a 5-point likert-type scale from 0 = *not at all* to 4 = *extremely*. The HRS-II has two reverse coded items. The first reverse coded item (item 3; *difficulty*), requests raters to indicate how difficult the homework activity was for the client. The original codes ranged from 0 (*not at all*) to 4 (*extremely*). The second item that is reverse coded is item 4 (*obstacles*), with anchors of 0 (not at all) to 4 (extensive). These two items reflect the level of homework non-completion. When reverse coded, ratings of 4 indicate a homework experience free of difficulty or obstacles. Reverse coding of these two items controls for rating bias as well as ensuring that items reflecting noncompletion are deducted from the overall score. All other HRS-II items represent aspects of homework relating to increased completion. A copy of the HRS-II is included in Appendix D.

The anchors are comprised of a number of adverbial qualifiers, such as *not at all*, *somewhat*, *moderately*, *very* and *extremely*. The choice of adverb qualifiers was made on providing an equal psychological distance between points on the scale (Kazantzis, Deane, & Ronan, 2005). Research into the metric characteristic of adverb qualifiers indicate that the anchors used in this measure *do* define rating positions that are approximately equidistantly spaced (Cliff, 1959; Howe, 1966).

Assignment Compliance Rating Scale (ACRS; Primakoff, Epstein, & Covi, 1986).

This therapist-rated, single-item measure was designed to assess the extent to which a client has complied with homework assignment(s) from the preceding session. The ratings range from 0 (the client did not attempt the assignment) to 6 (the client did more of the assignment than was requested). This measure reflects solely the quantity of homework completed, but not the quality of completion. However, this measure has been identified as one of the few homework measures used repeatedly in research (Kazantzis, Deane & Ronan, 2005), and is included here for comparison. To date, there has been no psychometric evaluation published for the ACRS (Whittal, Thordarson, & McLean, 2005). The ACRS is included in Appendix E.

Procedure

As with Study One, the observers supplied ratings once finished viewing each session on DVD. For there to be homework completion ratings, there needed to be homework reviewed in session. Ratings were made based on the review of homework and ensuing therapist and client behaviours throughout the session. The HRS-II was rated before the ACRS and following the rating of measures used in Study One.

Data Analysis Procedures

The data were treated similarly to the method employed in Study One (see Chapter 7). Analyses were made using SPSS for Windows, version 13 (SPSS Inc, 2004). Prior to analyis, the data were checked for missing values, and decisions were made regarding the treatment of these values. A discussion on the treatment of missing values is detailed in the results for Study Two (Chapter 11).

As noted above, the HRS-II had two items requiring reverse coding. All other data was keyed in as coded on the scales. These items were recoded prior to data analysis.

The statistical assumptions of the tests used in the present study were checked prior to conducting the main analyses. Where appropriate, SPSS for Windows was used to validate these assumptions. To ensure that the data in the sample met with the assumption of normal distribution inherent in the planned analyses, mean scores and standard deviations were generated for the HRS-II and ACRS data. As with Study One, univariate outliers were detected by first transforming the variables of interest to *z*-scores (Hair, Anderson, Tatham, & Black, 1998). Distributions were checked based on viewing the box and whisker plots, the histograms and results of Kolmogrov-Smirnov test of normality (Hair, Anderson, Tatham & Black, 1998).

Score Reliability

Estimates of interrater reliability were calculated for each of the two scales using intraclass correlation coefficient calculations. To complement the ICC estimates of reliability, two percentage scores were computed, one indicating absolute agreement, and one representing rater agreement based on scores within one point of each other. It has been suggested that rater reliability should be determined by a number of methods for the estimation of reliability and not based solely on one method, such as the ICC (Rankin & Stokes, 1998).

Correlational Analysis

While scores on the HRS-II were normally distributed, ACRS scores were markedly different from a normal distribution, as assessed by viewing box and whisker

plots and reviewing the results of the Kolmogrov-Sminov test. Logarithmic transformations did not improve the distribution of ACRS scores, so correlational analysis was made using Spearman's rho, a non-parametric equivalent to Pearson's *r*. The resulting correlation matrix was then checked for the degree of association between individual HRS-II items, the fullscale HRS-II total and the ACRS scores.

As a further assessment of the relationship between HRS-II and ACRS scores, the Bland-Altman procedure for determining the limits of agreement. This procedure is common in biomedical literature for the assessment of agreement between two clinical measures (for a review, see Dewitte, Fierens, Stock, & Thienpont, 2002), and has begun to be used in some clinical psychology research (Davidson, Obonsawin, Seils, & Patience, 2003). While correlational analyses provides evidence of a linear relationship between two measures, it has been argued that this may not be sufficient in determining the level of agreement between measures (Bland & Altman, 1986, 2003; Dewitte, Fierens, Stock, & Thienpont, 2002; Rankin & Stokes, 1998). By plotting the differences in scores against the average of the summed scores of the measure, a graphical representation of the limits of agreement between two measures is produced (Bland & Altman, 1986, 2003). From this bias in scoring and measure sensitivity can be determined. In the present study, analyses were made by first transforming total scores into percentage scores, as this would make interpretation easier (Dewitte, Fierens, Stock, & Thienpont, 2002).

Factor Analysis

It was decided that factor analysis would provide an initial indication of the construct validity of the HRS-II when used by independent raters. Before performing factor analysis, the data were checked for suitability based on the assumptions and

conventions of exploratory factor analysis. A correlation matrix of individual HRS-II items and the fullscale HRS-II was generated using Pearsons *r*. Correlations between items were inspected to determine whether a) most correlations exceded .30, and b) these correlations were significant (Hair, Anderson, Tatham, & Black, 1998). The Kaiser-Meyer-Olkin measure of sampling adequacy was assessed to determine whether it exceeded the recommended value of 0.6 (Dancey & Reidy, 2002; Kaiser, 1970, 1974).

Exploratory factor analysis using principal component factor analysis (PCA) was conducted using the data reduction function of SPSS 13 (SPSS, 2004). Exploratory factor analysis provides an initial indication of the underlying factor structure of a measure and is indicated when there are no specific expectations regarding the number, or nature of, the underlying constructs or factors (Thompson, 2004). PCA is the most commonly used when undertaking exploratory factor analyis (Hair, Anderson, Tatham, & Black, 1998; Kline, 1998; Thompson, 2004).

Decisions on the number of factors to include were made using two critera. The first step was the evaluation of eigenvalues based on the Kaiser eigenvalue rule, where factors are retained when the eigen-value exceeds 1 (Thompson, 2004). Secondly, an evaluation of the scree plot was used to determine if there was a clear levelling in the plot of results, which would indicate the point at which factor extraction be stopped Hair, Anderson, Tatham, & Black, 1998; Thompson, 2004).

The rationale for using orthogonal rotation over other forms of rotation was made for two reasons. The first was the ease in interpretation, which is made easier using orthogonal rotation, as the discussion of each factor can be made without regard to other factors (Hair, Anderson, Tatham, & Black, 1998; Streiner, 1994; Thompson, 2004). Secondly, and related, the use of a small sample, as was the case in this study,

can be problematic in interpretation (Hair, Anderson, Tatham & Black, 1998), therefore, using orthogonal rotation allowed for a simple model which can provide a starting point for future research.

Ideally, a cross-validation of this analysis would also have been performed by randomly splitting the data into two groups and comparing the results of a factor analysis with each half (Hair, Anderson, Tahtam, & Black, 1998), however, the sample obtained was not large enough to permit this.

Variability in Client Homework Completion

To assess differences in pattern of client completion of homework throughout the course of homework, means and standard deviations for each item of the HRS-II over the thre phases of therapy. The means for the HRS-II items were then graphically presented to indicate the temporal pattern of the behavioural and cognitive determinants of homework completion. That is, each item was graphically presented, over the course of therapy, with the other items representing the specific behavioural and cognitive domains described by Kazantzis and L'Abate (2005).

CHAPTER 11

Study Two Results

Overview

This chapter reports on the results of the study in relation to the second aim of examining the psychometric evidence for the HRS-II when being rated by independent observers. As part of a psychometric evaluation into the validity of the measure, seven specific hypotheses were posited. It was hypothesised that independent observers would be able to reliably rate the HRS-II. Further it was hypothesised that the HRS-II would be positively correlated with an existing measure of homework compliance, the ACRS. Similarly, it was hypothesised that the ACRS would show a strong, positive association with the HRS-II quantity item (item 2). It was also hypothesised that all other items of the HRS-II would show a positive relationship to the ACRS. As part of the assessment of the internal validity of the HRS-II, it was hypothesised that the underlying factor structure of the HRS-II would be consistent with the theoretical model of the behavioural and cognitive determinants of homework compliance proposed by Kazantzis and L'Abate (2005). Finally, it was hypothesised that there would be differences in the patterns of client homework completion across the phases of therapy.

Assessment of Missing Values

The sessions rated for Study Two were the same as those rated for Study One (the psychometric evaluation of the HAACS). Of the 74 rated sessions, there were 11 (14.8%) sessions identified by both raters as having no homework reviewed in session. As there was no homework reviewed, these sessions were excluded from analysis. Lack

of agreement between the two observers as to whether or not homework was reviewed resulted in a further 7 sessions (11%) being excluded from analysis on a listwise basis. The decision not to replace missing data was based on two considerations for the treatment of missing data. First, ratings were not recorded when a rater judged that no review of homework occurred. That is, when homework was not reviewed in session, a rating for homework review was not possible. Hair, Anderson , Tatham & Black (1998) refer to data missing in this type of circumstances, as censored data, a form of missing data termed ignorable. That is, the need for applying specific remedies for missing data is not required. In this instance, rater disagreement as to the identification of therapist homework review behaviours was not so much an issue of missing data, per se, but an issue of deciding whether to impute a value when rater disagreement occurred. The second consideration was the issue of imputing values when there is a large amount of missing data. Replacing with the mean artificially deflates the estimates of the standard deviation of the mean, resulting in a narrow band of confidence intervals (Streiner, 2002). The result of this is an increase in the possibility for Type I errors, according to Streiner. Based on these considerations, it was decided not to impute value when missing data occurred due to observer disagreement as to the occurrence of homework review in a viewed session. Ratings were made based on the judgement of experienced clinicians, and it was necessary to trust this judgement. For these reasons, only sessions where there was complete agreement as to the occurrence of homework review were included for analysis. One session (1.6%) of the ACRS was not rated due to rater error and was excluded listwise from analyses.

Session characteristics

There were 63 sessions where homework was reviewed in-session. The rated sample consisted of 6 therapists, with 21 clients represented in this sessions included for analysis. The mean number of sessions per client was 3 with a range of 1 through 5 sessions per client represented in this sample. As with the HAACS data set, one therapist provided almost 50 percent of the rated sessions (46%). Table 17 outlines the therapist distribution in this sample.

Table 17

Frequency distribution of therapist

Therapist	Frequency (n)	Percent
Therapist A	12	19.00
Therapist B	6	9.50
Therapist C	5	7.90
Therapist D	29	46.00
Therapist E	6	9.50
Therapist F	5	7.90
Total	63	100.00

Score Reliability

In order to assess the reliability of the HRS-II scoring between raters, two-way, random effects model of intraclass correlation coefficients were computed (see the Statistical Procedures in Chapter 7 for a description of the model). The overall ICC (2,2) was .82 , with 95% CI .79 to .85. This is within the range for excellent rater

agreement (Cichetti, 1994; Fleisss, 1981). ICC (2,2) were computed for each item of the HRS-II and the ACRS and are presented in Table 18.

Table 18

Intraclass Correlation Coefficients (ICC) for Items on the HRS-II and ACRS

Item	ICC	95% CI	
		Lower	Upper
Quantity	.79	.58	.88
Quality	.85	.73	.91
Difficulty	**.36**	-.06	.62
Obstacles	.63	.28	.80
Comprehension	.51	.16	.71
Rationale	**.20**	-.27	.51
Collaboration [a]			
Specificity [a]			
Match with goals	**.27**	-.22	.56
Pleasure	.74	.56	.85
Mastery	.83	.70	.90
Progress	.81	.68	.89
ACRS	.85	.73	.91

Note. HRS-II n = 56. ACRS n = 55. HRS-II = Homework Rating Scale-II. ACRS = Assignment Compliance Rating Scale. Items shown in boldface indicate poor rater agreement.

[a] *Collaboration* and *Specificity* items of HRS-II had zero variance in terms and ICC(2,2) could not be computed.

*p <.05, ** p <.01, *** p < .001.

As can be seen in Table 17, the ICC (2,2) range from .20 to .85, with three items having ICC rating below .40, suggesting that the rating for *match with therapy goals*, *rationale*, and *difficulty* items were more difficult for the two observers to agree on. It should be noted that *collaboration* and *specificity* had zero variance, which precluded the generation of ICC for these two items; variance is a necessary prerequisite for reliability (Nunnally & Bernstein, 1994), It has been noted that reliability estimates may be misleading, due to their intrinsic reliance upon either chance or variance (Hall, Groome, Streiner, & Rochon, in press; Looney, 2000; Thompson, 2003; Uebersax, 1987).

Rankin and Stokes (1998) suggest that ICC alone may not provide a true indication of reliability, and for that reason, complementary methods for assessing reliability should be employed. In order to complement the ICC calculations, two percentage agreement scores were calculated. The first reflects absolute agreement between raters, with the second percent agreement indicating widening of parameters to include scores within one point of each other. Absolute agreement percentage scores ranged from 35.71% for *difficulty* and *obstacles* to 96.40% for *collaboration* and *specificity*. It should be noted that a number of items having high ICC estimates showed considerably less agreement based on absolute agreement percentage scores. These are shown in bold in Table 19. Other items showed greater agreement using absolute agreement percentage scores (shown in italics in Table 19). When agreement was based on scores within 1 point of each other, all HRS-II items showed greater than 80% agreement. As an example, "*difficulty*" and "*obstacles*" had low ICC reliability estimates and absolute agreement percentages, but agreement rose to 82.14% when agreement was based on scores being within 1 point of each other. These results

suggest that, overall, raters applied similar ratings to the items, but occasionally differed as to the exact degree (hypothesis 13)

Table 19

Percent Agreement Between Raterson HRS-II Items Using Absolute and Within 1-Point Agreement

	Percent agreement	
Item	Absolute	Within 1 point
Quantity	**51.78**	89.29
Quality	**41.07**	94.64
Difficulty	35.71	82.14
Obstacles	35.71	82.14
Comprehension	42.86	83.92
Rationale	*50.00*	98.21
Collaboration	96.40	100.00
Specificity	96.40	100.00
Match with goals	*46.46*	80.36
Pleasure	**48.21**	92.86
Mastery	**55.36**	89.29
Progress	**48.21**	92.86
ACRS	56.36	94.55

Note. HRS-II items n= 56. ACRS n = 55. HRS-II = Homework Rating Scale-II. ACRS = Assignment Compliance Rating Scale.

Assessment of Normality

Prior to undertaking analyses, the data were checked for violations of the assumption of normality. There were no outliers detected for HRS-II data based on

box and whisker plots and stem and leaf plots. The Kolmogorov-Smirnov (Lilliefors correction) was not significant, $p>.20$, indicating that the HRS-II data did not differ from a normal distribution. The ACRS data produced both extreme outliers and a significant result on the Kolmogorov-Smirnov (Lilliefors Correction), $p < .001$. This indicated a distribution that was significantly different from normal. Transformation of the ACRS data did not produce a change in the distribution, so it was decided to use non-parametric analyses when using ACRS data in the subsequent correlational analyses.

Item and total means were computed for the HRS-II and ACRS with the results presented in Table 20. The results indicate that therapist *rationale*, *collaboration* and *specificity*, items 6, 7 and 8 respectively, showed less variance than other HRS-II items and ACRS as indicated by the minimum and maximum responses. The largest range in scores were for *quantity* and *quality* (HRS-II items 1 and 2). Overall, there was consistency between mean scores on the total homework completion measures (HRS-II and ACRS), indicating that there is satisfactory internal consistency and a degree of concurrent validity between the two measures.

Table 20

Mean Item and Scale totals showing Range, Mean, Standard Deviations and Variance for HRS-II and ACRS

					95% CI	
Item	Minimum	Maximum	M	SD	Lower	Upper
1. Quantity	0.00	4.00	2.46	1.00	2.20	2.73
2. Quality	0.00	4.00	2.25	1.10	1.96	2.54
3. Difficulty	1.50	4.00	2.58	0.74	2.39	2.79
4. Obstacles	1.00	4.00	2.69	0.89	2.45	2.93
5. Comprehension	0.50	4.00	2.69	0.76	2.49	2.90
6. Rationale	1.00	2.50	2.08	0.39	1.98	2.18
7. Collaboration	2.00	2.50	2.02	0.09	1.99	2.04
8. Specificity	1.00	2.50	1.99	0.15	1.95	2.03
9. Match with goals	1.50	3.50	2.78	0.43	2.66	2.89
10. Pleasure	0.00	3.00	1.67	0.84	0.95	1.39
11. Mastery	0.00	3.50	1.62	1.05	1.34	1.90
12. Progress	0.00	4.00	2.04	1.06	1.75	2.32
Mean total HRS	11.00	38.00	26.38	6.59	24.61	28.14
ACRS	1.00	5.50	4.20	1.06	3.91	4.49

Note. HRS-II n = 56. ACRS n = 55. HRS-II = Homework Rating Scale-II. ACRS = Assignment Compliance rating Scale.

Minimum score for all variables = 0. Maximum possible scores are as follows: individual HRS-II item = 4; fullscale HRS = 48; ACRS = 6.

Correlational Analysis

As part of the psychometric evaluation of the HRS-II, Spearman's *rho* correlation coefficients were computed using the mean HRS-II ratings and total HRS-II scores and ACRS scores. The inclusion of individual HRS-II items in this analysis was to assess whether or not the HRS-II items provided different information than the ACRS. Table 21 shows that, in general, the items show a positive association with each other, with most of the correlations over r_s = .35. Correlation between the

total scale HRS-II and ACRS was $r_s = .63$ ($p < .01$), indicating a moderate degree of association between both measures (hypothesis). While most correlations were positive, correlations with collaboration tended to be weak and negative (hypothesis). Of special interest is the moderate negative correlation between specificity and collaboration ($r_s = -.55$), indicating that one of these items was consistently rated lower than the other. Further, it should be noted that there are weak relationships between ACRS scores and scores on the process items of the HRS-II (i.e., *rationale*, *collaboration*, *specificity*, and *match with therapy goals*). The highest correlations between ACRS and HRS-II scores were with *quantity* ($r_s = .73$, $p < .01$) and *quality* ($r_s = .70$, $p < .01$), which were stronger than the correlations between ACRS and HRS total scale (hypothesis). The results indicate that there is a moderate degree of concurrent validity between the two measures. Further, these results indicate that while there are positive correlations between the HRS-II item scores and ACRS scores, the low explained variance between the full-scale HRS-II and ACRS ($r_s^2 = .40$) indicates that there are differences not accounted for by the scores alone.

Table 21

Correlation Matrix of HRS-II Items, HRS-II Total Score and ACRS Scores[a]

Measure	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1. Quantity	--													
2. Quality	.89**	--												
3. Difficulty	.47**	.55**	--											
4. Obstacles	.69**	.72**	.65**	--										
5. Comprehension	.72**	.74**	.50**	.52**	--									
6. Rationale	.28*	.34*	.31*	.39**	.35**	--								
7. Collaboration	-.03	-.13	-.03	.02	.06	.24	--							
8. Specificity	.13	.27	.17	.16	.00	.00	**-.51****	--						
9. Match	.27*	.34*	.20	.31*	.28*	.32*	-.03	.25	--					
10. Pleasure	.53**	.63**	.41**	.47**	.61**	.48**	-.10	.23	.42**	--				
11. Mastery	.56**	.71**	.38**	.54**	.52**	.45**	-.07	.24	.49**	.80**	--			
12. Progress	.61**	.76**	.45**	.57**	.59**	.42**	-.08	.29*	.51**	.79**	.93**	--		
13. HRS-II	.80**	.89**	.62**	.77**	.75**	.52**	-.03	.26	.52**	.81**	.88**	.90**	--	
14. ACRS	.73**	.70**	.47**	.58**	.59**	**.27***	**.03**	**.10**	**.18**	.46**	.48**	.49**	.63**	--

Note. n = 55. HRS-II = Homework Rating Scale-II. ACRS = Assignment Compliance Rating Scale. Correlations in boldface indicate correlations of special interst.

[a] Correlations using Spearman' *rho*

*p <.05; **p <.01,

To further assess the degree of agreement between the HRS-II and ACRS scores, the total scores of the HRS-II and ACRS were analysed using the Bland-Altman method for assessing agreement between two methods of clinical measurement (Bland & Altman, 1986). Although correlational analyses provides evidence of a linear relationship between two measures, it has been argued that this may not be sufficient in determining the level of agreement between measures (Bland & Altman, 1986,2003; Dewitte, Fierens, Stock, & Thienpont, 2002; Hawkins, 2002; Rankin & Stokes, 1998). By plotting the differences in scores against the average of the summed scores of the measure, a graphical representation of the agreement between the two measures was produced. As the two measures did not share the same metric, direct comparison by this method required transformation prior to conducting analysis. Although the ACRS score did not follow a normal distribution, this is not a prerequisite for determining the degree of agreement between two measures (Bland & Altman, 1986). Therefore, for ease of interpretation, scores for the two measures were transformed into percentage scores (i.e., scores out of 100).

As seen in the Bland-Altman plot, Figure 4, the mean difference between HRS-II and ACRS scores was -15.06 percentage points with 95% CI of the difference -17.75 to -12.55. The mean difference, or bias, between measures was -15.06, with the 95% CI for the bias ranging from -17.75 to -.12.25. The limits of agreement ranged from -38.54 to 8.42, indicating that 95% of difference scores fell between 8.42 above the mean difference to 38.54 below the mean difference, a fairly variable range in difference scores. The ACRS scores could therefore be 38.54% below, or 8.42% above the HRS-II scores. Plotting the difference against the mean, allows for the investigation of the relationship between error measurement and the true value (Davidson, Obonsawin, Seils, & Patience, 2003). Noticeable in Figure 4 are

two distinct clusters of scores indicating that, at low scores, the difference between the HRS-II and ACRS scores tends to be positive, indicating a bias towards HRS-II scores. At higher scores, the ACRS tends to be rated higher than the HRS-II, as indicated by the negative difference. The equations used in calculating Bland-Altman limits of agreement are included in Appendix G.

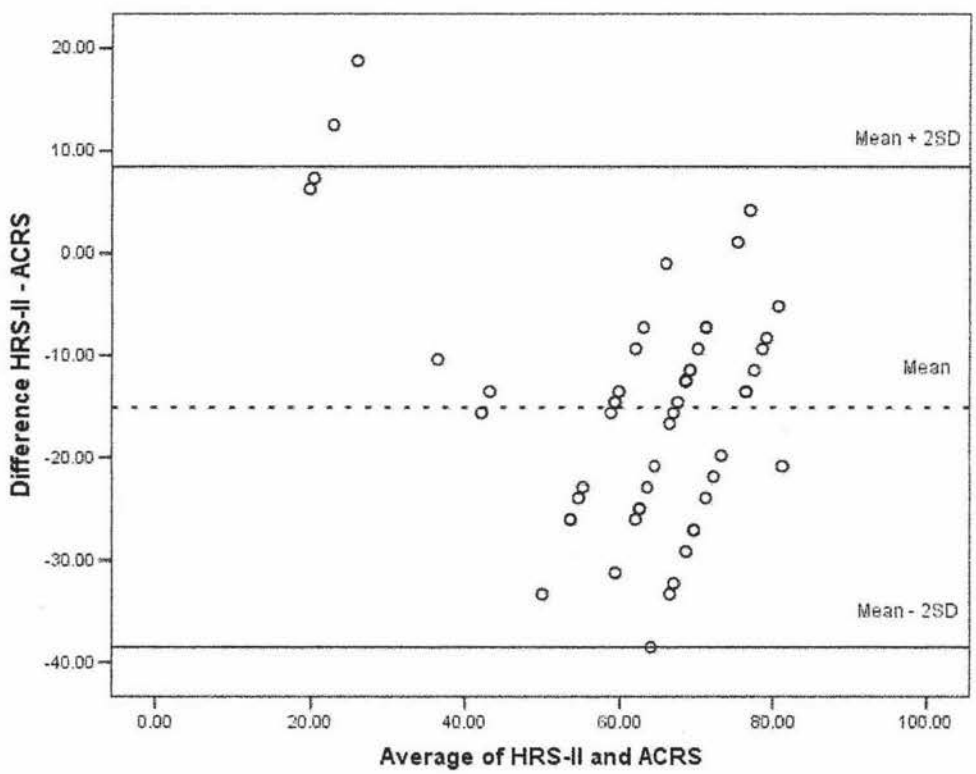

Figure 4. Bland–Altman plot for the difference between percent scores for the Homework Rating Scale-II and the Assignment Completion Rating Scale versus the magnitude of the percentage scores for the average of the two measures. The 2 standard deviations of the differences are shown to provide an indicationof the variability in difference scores. The short dashed line indicates the mean difference between the two measures, the solid lines indicate 2 standard deviation limits of agreement. See text for a detailed description.

Factor Analysis of HRS-II

As part of the psychometric analysis of the HRS-II, exploratory factor analysis was undertaken in order to derive the underlying dimensions that described the process of homework compliance in this sample. The authors of the measure envisage

five cognitive and behavioural determinants underlying the process of homework completion. These domains include beliefs, situation, behaviour, consequence and synthesis (Kazantzis & L'Abate, 2005; Kazantzis, Ronan & Deane, 2005). The items of the HRS-II were developed to assess specific behaviours associated with homework completion.

The twelve items on the HRS-II (therapist version) were subjected to Principal Component Analysis (PCA) in order determine whether or not factors based on independent observer ratings would be consistent with factors found in previous research using therapist and client ratings of homework completion. The overall goal of applying PCA was to identify and summarize patterns of correlations among the observed variables and to reduce these variables to a few, meaningful factors (Tabachnick & Fidell, 2001).

A total of 55 session ratings were analysed. Prior to analysis, there were a number of sample size conventions for conducting exploratory factor analysis to be considered. One convention is to have a sample as close to, or above 100 (Thompson, 2004). As a general rule, it is recommended that the minimum number of observations should be a ratio of five observations per variable to be analysed (Hare, Anderson, Tatham, & Black, 1998; Thompson, 2004). However, MacCallum, Widaman, Zhang and Hong (1999) have found that sample sizes as low as 60 can accurately reproduce population pattern coefficients, especially if communalities all exceed .60 or greater.

Prior to performing PCA, the data were assessed for its suitability for factor analysis. An inspection of the correlation matrix[3] revealed that most correlation

[3] Although a correlation matrix had been generated using Spearman's rho, a new correlation matrix using Pearson's product moment correlations was generated. This Correlation matrix can be found in Appendix H.

coefficients between HRS-II items as being equal to or above .3, the recommended minimum level for inclusion for factor analysis. (Hair, Anderson, Tatham, & Black, 1998). The Kaiser-Meyer-Olkin measure of sampling adequacy was 0.81, which exceeded the recommended value of 0.6 (Dancey & Reidy, 2002; Kaiser, 1970, 1974). Based on these underlying assumptions of factor analysis being met, it was decided that factor analysis was appropriate (Hair, Anderson, Tatham, & Black, 1998).

Initial analyses revealed the presence of three components with eigen values greater than 1, which explained approximately 50%, 15% and 9% of the total variance respectively. Although the scree plot did not reveal a clear break after the third variable, 73% of the total variance was explained by the first three components.

To aid in the interpretation of the components, varimax rotation was applied. The utility in an orthogonal rotation over an oblique rotation, such as direct oblimin, is that it provides no redundant information (Bryman & Cramer, 2005), therefore simplifying the interpretation That is, orthogonal rotation forces scores on factors to be unrelated, and therefore easier to interpret (Hare, Anderson, Tatham, & Black, 1998; Thompson, 2004).

The rotated solution is presented in Table 22, and shows three Factors, with Factor 1 accounting for 33% of variance; Factor 2 accounting for 26% of the variance and Factor 3, accounting for 15% of the variance. In total, 73% of variance in scores of the measure was explained by the three components. It should be noted that both the unrotated and rotated solutions explained the same amount of overall variance. Appendix I provides the details of the scree plot and the unrotated loadings.

Table 22

Summary of items and factor loading for Varimax Rotation Solution for HRS-II

HRS Items		Factor 1	Factor 2	Factor 3	Communality
Quality	(1)	**.85***			.90
Quantity	(2)	**.84***			.79
Obstacles	(4)	**.82***			.73
Difficulty	(3)	**.75***			.58
Comprehension	(5)	**.71**			.69
Mastery	(11)		**.77***		.81
Pleasure	(10)		**.75***		.73
Match with goals	(9)		**.75***		.57
Progress	(12)	**.55**	**.73**		.87
Rationale	(6)		**.61**		.54
Collaboration	(7)			**-.89***	.80
Specificity	(8)			**.87***	.79

Note. n=56. HRS-II= Homework Rating Scale-II.

* significance based on sample size, p = .05. 80% power and standard error assumed to be twice that of conventional correlations

According to Hair, Anderson, Tatham, and Black (1998), factor loadings greater than .50 are considered practically significant. Statistical significance, based on sample size and with $p \leq .05$, is determined by factor loadings greater than .75 (Hair, Anderson, Tatham, & Black, 1998). In this instance, all factors loadings achieved practical significance, and most achieved statistical significance. The communality coefficients, h^2, shown in Table 22 are the lower-bound estimates of score reliability (Thompson, 2004). That is, the value of the communality coefficient indicates that the

reliability of scores on a variable is no lower than the value of h^2, and may be higher (Thompson, 2003). In this instance, h^2 values ranged from .58 to .90, indicating that the variables provide sufficient explanation for the result (Hair, Anderson, Tatham, & Black, 1998; Thompson, 2004). The rotated factor solution shows the presence of a simple structure, with strong loadings for the components on each factor (Thompson, 2004).

Factor 1 consisted of six items: 1 (*quality*), 2 (*quantity*), 4 (*obstacles*), 3 (*difficulty*), and 5 (*comprehension*). The factor encompasses items from four of the domains described by Kazantzis and Deane (2005). Two items match items in the behaviour domain, as well as including one item each from the domains of consequences, situations, and beliefs. Thus, Factor 1 was labelled Antecedents and Homework Completion. It should be noted, however, that item 12 (*progress*) loaded onto this factor to a lesser extent (.55). The decision not to include this element in this factor was based on recommendations by Nunnally and Bernstein (1994) that the inclusion of elements with factor loadings between .30 and .60 may be misleading. Further, according to Kim and Mueller (1978), the decision to include or exclude variables may be based on underlying meaning rather than factor loadings. In this instance, it was decided that including progress would be contrary to the rationale for using an orthogonal rotation; that is, the reduction of elements into a simple, interpretable solution.

Factor 2 consisted of five items, 11 (mastery), 10 (pleasure), 9 (match with goals), 12 (progress) and 6 (rationale). These items matched items in the synthesis and consequences domain, with one item from the beliefs domain. Thus Factor 2 was labelled as Synthesis and Consequences.

Factor 3 consisted of two items, 7 (collaboration) and 8 (specificity). These items are derived from the beliefs domain, therefore this factor was labelled Beliefs. However, this factor was comprised of only two variables, less than the recommended minimum of three variables (Streiner, 1994). Thus, this factor should be ignored, leaving a simple two-factor solution for the HRS-II for this sample.

Cronbach's alpha was calculated for each of the two remaining factors to determine the reliability of the factors. Generally, an alpha value greater than .7 is considered acceptable. Both Factor 1 and Factor 2 exceed that limit with alphas of .90 and .86 respectively.

Comparison of Factor Analysis Results

Results from the factor analysis in the present study were consistent with the five cognitive and behavioural determinants of homework completion presented in Kazantzis and L'Abate (2005) (hypothesis 18). One aim of the present study was to compare factor analysis results for ratings made by independent observers to results from another preliminary study using therapist and client ratings of homework completion. It should be noted that this comparison is only a preliminary evaluation and solely based upon qualitative analysis rather than using a statistical approach such as the Procrustes procedure (for a description of the procedure, see Hill, O'Grady, & Elkin, 1992), as the data were not available for analysis. Bjornholdt (2006) studied differences in therapist and client ratings of homework completion using the HRS-II therapist and client versions. Principal component analyses using oblique rotation produced similar factors for both client and therapist. In order to compare the findings, the findings from this study were tabulated with the results from the Bjornholdt study and are shown in Table 23.

Table 23

Comparison of Factor Analyses based on Independent Observer, Therapist and Client Ratings of HRS-II.

	Source of Ratings		
	Independent Observer	Therapist	Client
Factor 1	(Antecedents and Completion)	(Benefit and Completion)	(Benefit and Completion)
	Quality	Quality	Quality
	Quantity	Quantity	Quantity
	Obstacles	Mastery	Mastery
	Difficulty	Pleasure	Pleasure
	Comprehension	Progress	Progress
Factor 2	(Synthesis and Consequences)	(Beliefs)	(Cost and Completion
	Mastery	Comprehension	Difficulty
	Pleasure	Match	Obstacles
	Match	Rationale	Quantity
	Progress	Collaboration	Quality
	Rationale	Specificity	
Factor 3			(Beliefs)
	Collaboration	*Difficulty*	Comprehension
	Specificity		Match
			Rationale
			Collaboration
			Specificity

Note. HRS-II = Homework Rating Scale-II. Therapist and Client factors come from Bjornholdt (2006) and are based on Principal Components Analysis using oblique rotation. Independent observer factors come from the present study.Labels for factors are shown in brackets. Variables shown in italics indicate variables not included, as the number of variables for the corresponding factor were less than three.

Therapist and client ratings of homework completion shared variables on two factors (client Factor 1 and therapist Factor 1; client Factor 3 and therapist Factor 2), but the amount of explained variance differed, suggesting the relative importance of these variables may have differed between client and therapist ratings (Thompson, 2004). There were no shared factors between independent observer, therapist and client ratings, although Factor 1 for independent observer ratings and client Factor 2 were similar.

The amount of variance explained by client Factors one, two and three were 42.6 %, 13.4% and 10.8% respectively with a cumulative total of 66.8%. The amount of variance explained by therapist Factors one, two and three respectively were 38.3%, 14.3% and 10.3% with a total of 63.0%. The amount of variance explained by independent observer ratings in the present study for Factors one, two and three respectively were 33% 26% and 15% respectively, with a cumulative total of 73%.

Results suggested that *quantity* and *quality*, as rated by therapist and client, are primarily associated with the perceived benefits of homework completion (*pleasure*, *mastery* and *progress*), whereas independent observer ratings indicated an association between potential mediators and moderators of *quantity* and *quality*. That is, independent observer ratings indicated an association with variables present prior to undertaking homework, whereas therapist and client rarings indicated an association with variables subsequent to undertaking homework.

Difficulty and *obstacles* did not load sufficiently on any factor for therapist ratings, but were rated with higher relative importance in independent observer ratings than when rated by clients. The results showed an underlying association between difficulty, obstacles, quantity and quality for both client and independent observer ratings of homework completion. However, this association was not as strong

in client ratings where quantity and quality showed loadings of .47 and .45 respectively.

Variability in Client Homework Completion

One of the aims of this study was to examine differences in ratings of client homework completion across the phases of therapy. It was not expected that client completion would be consistent from phase to phase, as the therapeutic focus in CBT differs across the therapy, and the differences may impact on client involvement with their homework. Thus, it had been hypothesised that there would be differences in the pattern of client homework completion across the phases of therapy (hypothesis 19).

Prior to analysis, means and standard deviations were computed for the individual HRS-II items for each of the three phases of therapy (early, middle and late). As indicated in Table 24, most homework completion ratings did not differ substantially over the three phases of therapy. There were a few minor exceptions to this, with *quality*, *pleasure* and *progress* showing the slightly larger increases in ratings across therapy. Most ratings increased over therapy, with four exceptions: ratings of *difficulty* fell across all phases; *match with therapy goals* and ACRS rating fell slightly in the third phase of therapy. *Difficulty* is reverse coded, so lower scores represent an increase in difficulty, suggesting that clients experienced slightly more difficulty with homework in the middle and late phases in therapy.

Table 24

Means and Standard Deviations of Client Homework Completion Across Therapy Phases

	Phase of Therapy					
	Early		Middle		Late	
Variable	M	SD	M	SD	M	SD
Quantity	2.34	0.96	2.55	0.99	2.58	0.98
Quality	1.82	0.95	2.36	1.05	2.69	1.03
Difficulty	2.63	0.57	2.52	0.78	2.50	0.94
Obstacles	2.55	0.88	2.67	1.02	2.81	0.72
Comprehension	2.41	0.82	2.73	0.75	2.96	0.66
Rationale	2.03	0.42	2.05	0.42	2.19	0.33
Collaboration	2.05	0.16	2.00	0.00	2.00	0.00
Specificity	1.95	0.23	2.00	0.00	2.04	0.14
Match	2.68	0.34	2.86	0.36	2.69	0.60
Pleasure	0.89	0.74	1.21	0.78	1.46	0.97
Mastery	1.11	0.91	1.74	0.97	2.04	0.97
Progress	1.50	0.82	2.14	0.88	2.58	1.10
HRS-II	23.97	5.71	26.83	6.47	28.54	6.80
ACRS	4.21	1.06	4.24	0.89	4.21	1.03

Note. Early Phase n = 19. Middle Phase n = 21. Late Phase n = 12. Early Phase includes sessions 2-7. Middle Phase includes sessions 8-13. Late Phase includes sessions 14-20. Maximum scores for HRS-II = 48; Maximum score ACRS = 6; all other variables have a maximum score = 4. Minimum score for all variables = 0. The higher the score, the greater the homework completion.

In order to illustrate the temporal pattern of client homework completion, ratings of the HRS-II items were graphed according to the five behavioural and cognitive determinants of homework completion that underly the measure (i.e., beliefs, situation, behaviour, consequence and synthesis). To aid in the visual interpretation, *difficulty* and *obstacles* are reverse coded, so high values reflect higher levels of client experiencing of *difficulty* and *obstacles*, that is, less homework completion.

Client beliefs about homework suggest that client comprehension was consistently rated higher by the independent raters than were the other aspects of client beliefs (Figure 5). Ratings for collaboration fell slightly across the phases of therapy, with all other ratings of client beliefs rose. Figure 6 shows that ratings of obstacles met by the client fell throughout the course of therapy. Of particular interest are the ratings of the behaviour variables shown in Figure 7, which indicated that although the quantity of homework completed was rated highly throughout, there was a definite levelling off of the amount completed by the client between the middle and late phases of therapy. Quality, however, rose throughout the phases of therapy, ending with a higher rating than did quantity. Figure 8 shows the ratings for consequence items, which indicated that while ratings of client experiencing of difficulty remained fairly constant throughout therapy, ratings for pleasure and mastery increased. This suggested that the effects of difficulty was possibly counterbalanced, or moderated by other aspects of homework completion. Ratings for synthesis items revealed that client progress rose steadily throughout therapy, whereas ratings of the match with therapy goals peaked during the middle phase of therapy, as shown in Figure 9.

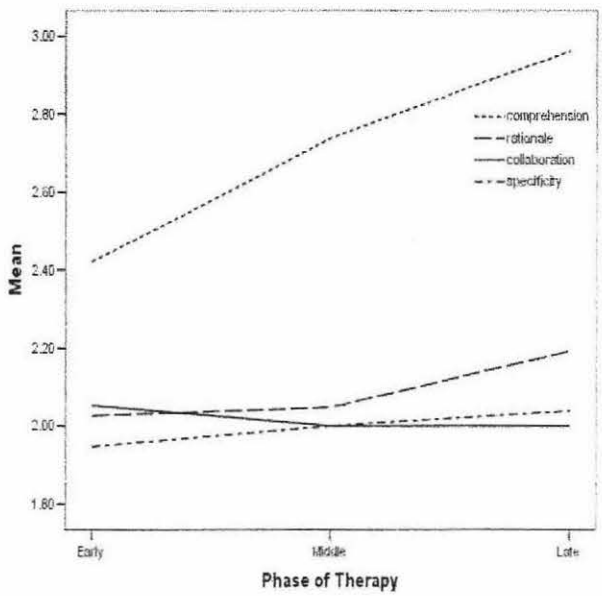

Figure 5. Independent observer ratings of client homework beliefs across phases of therapy.

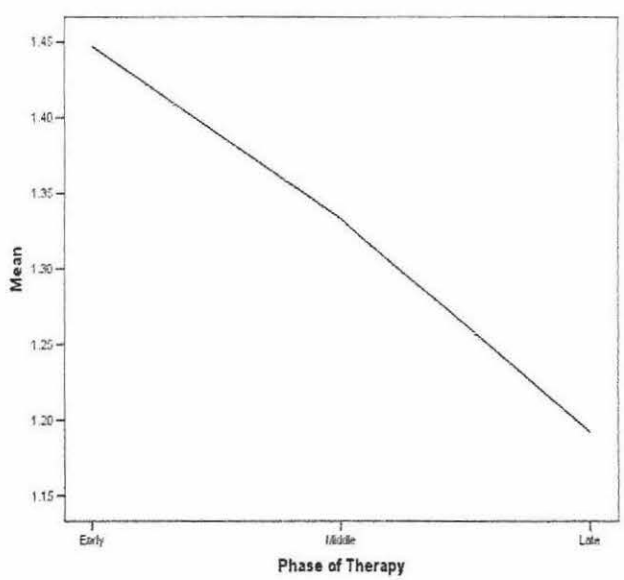

Figure 6. Independent observer raings of client homework obstacles across phases of therapy.

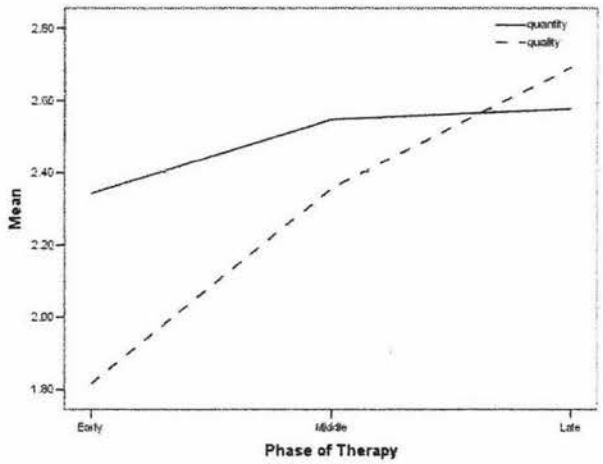

Figure 7. Independent observer ratings of client homework behaviours across phases of therapy.

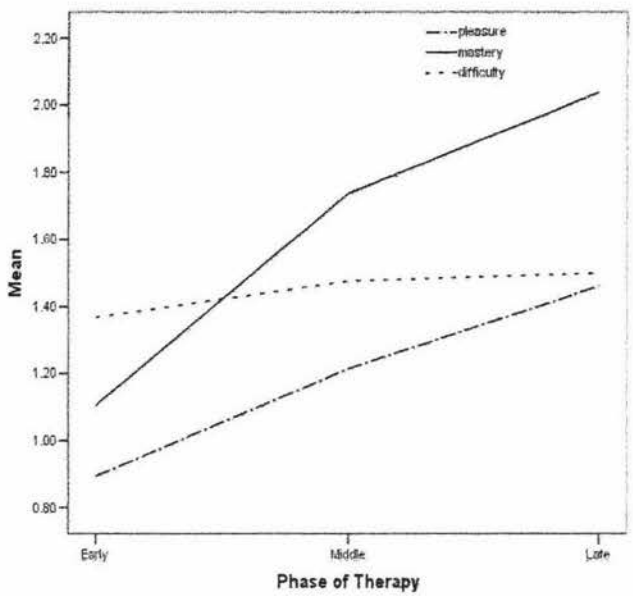

Figure 8. Independent observer ratings of client homework consequences across phases of therapy.

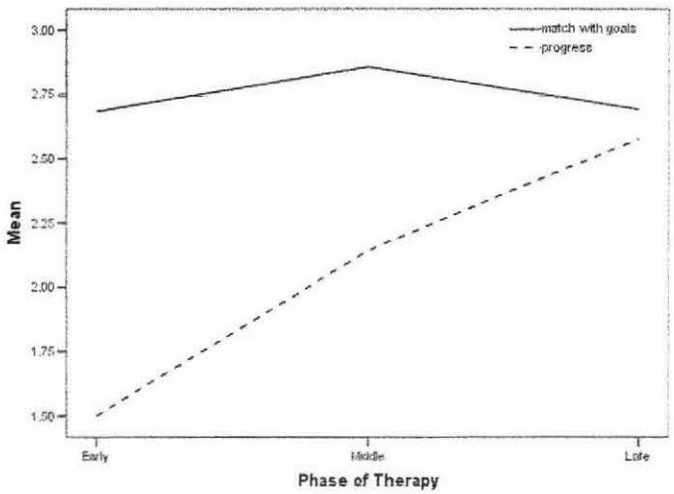

Figure 9. Independent observer ratings of client homework synthesis across phases of therapy.

CHAPTER 12

Discussion Study Two

Overview

As part of the ongoing process of gathering psychometric evidence for the HRS-II, one of the aims of this study was to determine whether or not independent raters could score the measure reliably. It has been noted that measures of homework completion tend to be subjective measures, either using client self-report or therapist reports. The use of independent observer reports may then aid in reducing bias inherent in therapist and client reports (Perepletchikova & Kazdin, 2005). Further, Doss (1999) suggests that measures of client change mechanisms need to be assessed directly.

As part of the ongoing process of gathering psychometric evidence for the HRS-II, one of the aims of this study was to determine whether or not independent raters could score the measure reliably. The utility in having observers rate in-session behaviours has been highlighted by many, including Perepletchikova and Kazdin (2005) and Elkin, et al. (2004). These authors recommend the use of independent raters to reduce the influence of therapist or client bias on results.

This chapter presents an overview of the main findings of the present study, with the results of each hypothesis discussed. The study's limitations as well as suggestions for future research and the clinical implications resulting from the present study will be discussed in Chapter 13, General Discussion and Conclusions.

Score Reliability

Overall, the interrater agreement for the measure indicated an excellent level of agreement (Fleiss, 1981) with an ICC of .82 for the total HRS-II. Most items of the HRS-II achieved a level of agreement using ICC over .60, which is acceptable and consistent with ICC calculations at the item level from other validation studies (e.g., Barber, Liese, & Abrams, 2003). However, *collaboration* and *specificity* showed no variance in scores, so ICC could not be computed. One potential explanation is that, for this sample, *collaboration* and *specificity* did not differ from throughout the rated sessions. However, this is unlikely, and the ratings for these items may actually indicate that *collaboration* and *specificity* items were more difficult to rate in this context for the raters. That is, it may have been difficult for raters to provide an accurate judgement of prior *collaboration* and *specificity* based on how these occurred in the rated session.

When estimates of reliability were based on less stringent criteria (agreement based on ratings being within 1 point of each other), agreement rose to over 80% for all items of the HRS-II. This suggested that using absolute agreement estimates for reliability may produce lower estimates of agreement when used in small samples due to the inherent effect of small variations on overall reliability (Hall, Groome, Streiner, & Rochon, in press; Looney, 2000; Thompson, 2003; Uebersax, 1987).

In all, although ICC estimates of reliability produced some mixed results at the item level, the results provide initial support for the hypothesis that independent raters would be able to reliably rate the HRS-II (hypothesis 13).

Correlational Analysis

In order to provide concurrent validity evidence, it was hypothesised that total HRS-II scores would be positively correlated to ACRS scores, a measure of client homework compliance presently used in research (hypothesis 14). The results showed a moderate association between the ACRS and total HRS-II scores, with 37% of shared variance. This indicated that the HRS-II and ACRS share some similarity in the assessment homework completion, but that there is a great deal of difference between the two measures not accounted for by the scores alone.

Given that the ACRS measures quantity of homework completion, it was hypothesised that there would be a strong, positive association between ACRS scores and scores for item 2 (quantity) of the HRS-II (hypothesis 15). This hypothesis was supported, with the HTS-II item *quantity* producing the highest correlation with the ACRS. In relation to the other items of the HRS-II, it was hypothesised that the individual HRS-II item scores would be positively correlated to each other (hypothesis 16). The results proved interesting with correlations between *collaboration* and the other HRS-II variables resulting in weak, negative associations. The exception to this was the significant moderate negative correlation produced between *collaboration* and *specificity*. This indicates the possible existence of a systematic bias in ratings, with *collaboration* being consistently rated higher than *specificity*, as indicated by the mean scores displayed in Table 20. There are a number of possible interpretations for this finding. First, the results could indicate that in the rated sessions therapists tended to be more *collaborative* than they were *specific*. That is, when there was *collaboration* in the designing of homework, there was a tendency for therapists to be less *specific* in the assigning of homework. However, ratings for *collaboration* and *specificity* required the independent observers to make clinical judgements

as to the degree of collaboration and specificity incurred in the previous session. Ratings may then not be as accurate in reflecting the true extent of *collaboration* and *specificity* that did occur, and may be more indicative of the rated session rather than what occurred in the previous session.

It had also been hypothesised that the individual HRS-II item scores would be positively correlated to the ACRS scores (hypothesis 17). Correlations between scores for individual items of the HRS-II and ACRS scores indicated significant moderate to large associations. The exceptions to this trend were associations between ACRS and specific process items of the HRS-II, where the relationship was minimal. This indicated that the ACRS and HRS-II item *rationale*, *collaboration*, *specificity* and *match with therapy goals* measure different aspects of homework completion than covered by the ACRS.

Given the results of the correlational analysis, there is support for the HRS-II as a measure of homework completion. The results indicated both concurrent and diverent validity between the HRS-II and ACRS measures. By this, the HRS-II measures homework completion, as doe the ACRS, but the two measures differ what they measure. Specifically, the HRS-II adds a depth and breadth to the measurement of homework completion not currently measured by the ACRS. There are a number of factors that would recommend the use of the HRS-II over the ACRS. While the ACRS is a quick and easy measure to use, there are inherent problems associated with the use of one-item measures, specifically reliability and discriminability (Kline, 1986). Further, internal consistency is directly related to the length of a measure, and coefficient alphas do not give a great indication of internal consistency when the number of items is small (Kline, 1986). For these reasons alone, the use of a single item measure to assess homework

completion cannot give as good an indication of the necessary components for improved outcome, which is the focus of most research in the psychotherapeutic use of homework.

Further, the results of the Bland-Altman procedure suggested that while there is a strong linear relationship between the ACRS and HRS-II scores, there appears to be differences in how the measures actually score homework completion. In the main ACRS percentage ratings of homework completion tended to be higher than percentage ratings for the HRS-II. Further, there was a large amount in variability in the range of ACRS percentage scores resulting in ACRS scores being between 8.42% above HRS-II scores to 38.54% below HRS-II scores. These findings suggested that, overall, the ACRS may not be as reliable in detecting differences in client homework completion as the HRS-II.

Factor Analysis

It was hypothesised (hypothesis 18) that factor analysis of the HRS-II would result in a simple, interpretable solution consistent with the cognitive and behavioural determinants of homework completion described by Kazantzis and L'Abate (2005). The results of the Principal Components factor analysis indicated the presence of two factors associated with independent observer ratings of therapist adherence behaviours in the integration of homework into practice. A third factor, comprised of collaboration and specificity was also indicated, but did not meet the minimum number of variables required for inclusion as a factor (Streiner, 1994; Thompson, 2004).

For this sample, the first factor was best described as Antecedents and Homework Completion. 1 (*quality*), 2 (*quantity*), 4 (*obstacles*), 3 (*difficulty*), and 5 (*comprehension*). Four of the five behavioural and cognitive domains for homework completion were included in this factor, indicating that the components represent the overall experience of homework

from designing homework through to the review of the experience. The results suggest that, as assessed by independent raters, the actual doing of the homework is influenced by the presence of obstacles and the client's perceived costs related to engaging with the homework. This result suggests similarity to the Theory of Planned Behaviour (Ajzen, 1991) in that behavioural enactment is related to the overall appraisal, or cost, of the behaviour.

Factor 2 was labelled Synthesis and Consequences. This factor suggested that client homework completion results in specific inferences made by the client about the utility of having complied with the homework task. That is, as a result of complying with homework, specific beliefs about the actual benefit from that compliance are formed. It is possible that this is similar to Festinger's theory of Cognitive Dissonance (1957), whereby behaviour enactment causes a shift in attitude towards the behaviour.

The underlying factor structure of the HRS-II, as determined from the ratings of independent observers suggests that independent raters may view homework completion differently than therapists and clients. In a recent study of therapist and client ratings of homework completion, Bjornholdt (2006) found similarities between client and therapist ratings of homework completion, with quality and quantity of homework associated with pleasure, mastery and progress. In the present study, quantity and quality were associated with difficulty, obstacles, and to a lesser degree, progress. This may indicate that therapist and client ratings of homework share a more positive focus, perhaps due to social desireabilty or the halo effect. Or, it may be that independent observers tend to focus on the negatives, or obstacles, to the completion of homework.

Variability in Client Homework Completion

Specific statistical analysis of client variability in homework completion was not possible in the present study for a number of unforseen reasons. First, the sample size was smaller than anticipated, as homework was reviewed in fewer sessions than originally anticipated. Further, there were insufficient sessions for each client across the phases of therapy to analyse client variability. However, based on the literature searches conducted and discussed in earlier chapters, the present study appears to be the first to plot the temporal pattern of homework completion variables. The vast majority of prior research on homework has tended to focus solely on either the quantity or quality of homework completion, while neglecting to examine other components of homework completion, The present study found that the temporal pattern of quantity and quality differs across the phases of therapy, suggesting that the measurement of one aspect may not necessarily be indicative of the other. That is, the evaluation of homework completion should not be limited to quantity or quality alone.

Given that the therapeutic focus changes throughout CBT (J. S. Beck, 1994), it was hypothesised that there would be differences in the pattern of client homework completion across the phases of therapy (hypothesis 19). The results indicated that ratings of the experiencing of *obstacles* by clients decreased considerably over time. Some variables remained fairly constant, such as *rationale*, *collaboration*, *specificity* and *difficulty*. Other variables showed relatively steady increases across the phases of therapy, such as *progress*, *mastery*, *comprehension*, *quality* and *pleasure*. Further still, the *quantity* of homework completed by clients showed a tendency to peak during the middle phase of therapy, with a distinct drop across the late phase of therapy.

The finding that the pattern for each variable differed from other variables is a potentially important. The pattern of ratings shown in the graphs indicated that the process of client homework completion resulted in different changes at different times for the clients. There is the possibility that the patterns may indicate differences due to therapist effects rather than by homework completion alone, as this potential confound was not controlled for. Further, as the graphs did not contain the same clients across all the phases of therapy, the patterns may be reflecting differences between clients not measured by the HRS-II, such as differences in client symptomatology.

CHAPTER 13

General Discussion and Conclusions

This study set out to provide psychometric evidence regarding two new measures, the HAACS and HRS-II. Chapter 9 (Study One) and Chapter 12 (Study Two) provided discussions detailing the results of the specific hypotheses of the present study. This final discussion is not intended to repeat or summarise the discussions from the previous chapters, but will discuss the limitations of the present study overall. Further, recommendations for future research will be discussed, as will the implications for clinical practice.

Limitations

Studies of therapist adherence and competence are generally limited by cost (Wampold & Bhati, 2004). As with any study, there were a number of limitations in the present study that need to be addressed. First, the small resulting sample size was both unexpected and unavoidable, as the data were provided from a previous study. Although 100 rateable sessions had been expected, there were a number of sessions that did not appear on the DVDs, and a large number of rateable sessions did not include a review of homework. Further, only sessions where the independent raters agreed on the occurrence of homework review were included for analysis. This resulted in only 63 sessions being included for analysis of the HRS-II. As such, the results of the present study should be considered preliminary and require replication with a larger sample. Specifically, the small sample size has impacted on the generaliseability of the findings for the factor analysis performed in Study Two (Hair, Anderson Tatham, & Black, 1998).

A further limitation that arose from the archived data was the unequal representation of both therapists and clients. In addition, the number of clients per therapist varied and did not allow for analyses to control for the effect of therapist-client dyads. Similarly, there were insufficient therapist-client pairings across the phases of therapy to address the possible confound of client factors on the findings in relation to the temporal nature of therapist adherence and competence. While there were 21 clients in the sample, there were few clients with sessions that spanned all three phases of therapy. This precluded the specific analysis of client variation in homework completion, as there needs to be two points of data to show change, and a minimum of three available data points to indicate a pattern in differences, as suggested by Bryk and Raudenbusch (1992).

While research into therapist effects on outcome has been mixed (for an overview of the research, see Chrits-Christoph & Gallup, 2006), therapist variability in both adherence and competence is an area that requires further inquiry. Results from this study supports the hypothesis that there would be an effect for therapist on adherence and competence, However, the presence of a large main effect for therapist may be due to differences in therapist distribution rather than indicating core differences. Elkin et al. (2005) point out that the use of small samples may inflated effect size when assessing therapist differences. Further to this, the combining of therapists from two different treatment protocols may be implicated in the finding of a main effect for therapist. That is, the results may indicate differences in therapists based on the specific therapy provided, rather than differences due only to individual therapists. Therefore, the conclusions from the present study should only be considered preliminary.

There were few sequential sessions available for rating. This may have impacted on the ratings of some items of the HRS-II requiring some contextual knowledge of the previous session, which would be available for client and therapist ratings of homework completion, but not necessarily to independent raters. Specifically, items such as *collaboration* and *specificity* refer to behaviours occurring in the previous session that may have influenced the completion of homework by a client. Consequently, there is some concern about the viability of rating these specific items without access to the previous session.

Further, the present study did not have access to client outcome measures, so no conclusions have been possible regarding the importance of homework completion on treatment outcome. As well, the results of the present study may have been confounded by client difference in symptom severity and rate of remission, which could have been controlled for had the outcome assessments been available. Similarly, as with Study One, there was unequal representation of therapists in this sample, and the results may indicate specific effects of the therapist rather than from homework completion. Had there been an equal representation and considerably more rateable sessions, therapist and client effects could have been controlled for, as well as providing an analysis between degree of homework completion and therapist adherence and competence in homework integration.

The lack of access to the data set from the Bjornholdt (2006) study precluded the direct comparison of the factor analyses using Procrustes procedure, which would have statistically tested the observer, therapist and client models of homework completion. As such, the comparison of independent observer, therapist and client factors is preliminary and should be used as a starting point for future comparisons.

While there are methodological limitations in this study related to the use of archived data, these limitations also add to the relative strengths of the study. The present study was, in essence, an observational study of a treatment protocol (see Dimidjian, et al., in press) not originally designed to assess adherence and compliance in the administration of homework. As such, the study does not suffer from the inherent biases that could exist had the therapists known they would be assessed on homework-related behaviours.

Recommendations for Future Research

The results of the present study provide preliminary evidence of the utility for the use of both the HAACS and HRS-II in research when rated by independent observers. It is expected that future research be undertaken to provide validation for the initial findings presented in this research. The use of larger samples and specific homework related treatment protocols will be able to assess the utility of these measures in a more appropriate context and may overcome some of the limitations found in the present study.

It is recommended that ratings of sequential sessions be used, in order to overcome potential difficulties for independent raters in rating certain aspects measured by the HRS-II. That is the possibility of raters having difficulty in providing an accurate assessment of *collaboration* and *specificity* items, which was indicated by the findings in the present study. Used in conjunction with outcome measures, this may help in the development of a model of client change processes throughout the course of therapy. The further assessment of the cognitive and behavioural determinants of homework completion has been indicated by the present research, as there appears to be differences

across the course of therapy that indicate that change in patterns of homework completion may occur at different times.

The use of the HAACS in conjunction with the HRS-II may provide an indication as to which therapist behaviours aid client homework completion and, possibly, when these behaviours are important in effecting change in the client. Thus, it is recommended that future research be designed to address this. That is, future research could be conducted to include the assessment of both therapist and client homework behaviours across the course of therapy. Studying the relationship between client homework compliance and therapist adherence and competence in using homework may provide useful information in predicting treatment outcome.

The present study has used a new scoring method for determining competence scores which has resulted in the apparent separation of adherence from competence, a problem in research highlighted as a confound in outcome studies (Crits-Critsoph & Mintz, 1991). It is recommended that future research test the new scoring method to assess its viability as an alternative scoring method.

Clinical Implications

There is potential for the HAACS and the HRS-II measures to be used in clinical practice. Specifically, it is expected that the HAACS may prove useful in training and ongoing supervision of CBT therapists, while the HRS-II may have utility in tracking client progress as well as possibly ascertaining information regarding client homework completion which may not be assessable within the dyad context.

As a method for assessing therapist adherence and competence in administering homework, it is anticipated that the HAACS will contribute to the field of clinical psychology in both the training and ongoing supervision of CBT therapists. Coupled with the guiding model for practice developed by Kazantzis and L'Abate (2005), the HAACS provides a theoretically and empirically grounded measure of therapist adherence and competence in homework administration. As an ongoing training and supervision tool, the HAACS indices provide a framework for the feedback of results. At an item level, the HAACS can be used to determine a therapist's relative strengths and weaknesses in the integration of homework into therapy. Although the measure is relatively long, and this may impact on the practicality if its use to rate every session, the HAACS may prove useful when used occasionally to identify specific areas for improvement in homework integration.

The HRS-II has been shown to have an acceptable degree of reliability when rated by independent observers. Data from the present study were found to provide support for the proposed cognitive and behavioural determinants of client homework completion (Kazantzis & L'Abate, 2005). However, due to the limitations of the present study, it was not possible to ascertain the relationship between homework completion variables and outcome. It could be argued that given the empirical and theoretical support for the use of homework, the use of a comprehensive measure of client homework completion may provide an important clinical tool useful in tracking individual client progress throughout the course of therapy.

Further, the use of independent raters in the assessment of homework completion may provide clinicians a more unbiased account of client homework completion. The comparison of HRS-II factors for client, therapist and independent observers indicated

that client and therapist rate homework completion similarly in relation to the quality and quantity of homework while independent observers tend to rate homework completion differently and tend to associate quantity and quality with obstacles and difficulty related to homework completion. Thus, the use of independent observer ratings of homework completion in clinical practice may provide an unbiased account of the actual degree of barriers experienced by the client. The importance of the identification of barriers to homework completion is an important aspect of the successful integration of homework into practice and has been highlighted as a key recommendation for therapist homework behaviours in promoting client engagement with homework (J. S. Beck, 1995; Kazantzis & L'Abate, 2005).

Conclusions

The present study provides initial psychometric evidence in support for the HAACS and HRS-II as viable measures for assessing client and therapist homework behaviours in CBT when assessed using independent observers. Both measures have firm empirical and theoretical groundings, and it is expected that the measures will undergo ongoing validation using empirical methods in order to gather further psychometric evidence

As a measure of therapist adherence and competence in the integration of homework into therapy, the present study has provided initial psychometric evidence of its utility when rated by independent raters. The HAACS is comprehensive and specific in its content, which addresses some concerns posited by Whisman (1993) and Shaw, et al. (1999) that therapist competence measures such as the CTS may be inadequate for the assessment of therapist competence. Further, the construction of the HAACS provides a

comprehensive description of what homework competence entails, an important step towards the understanding of the role of homework in therapy.

The present study provides initial psychometric evidence for the HRS-II, when rated by independent observers. While the ACRS provides a quick measure of the quantity of client homework completion, the HRS-II has been shown to assess different dimensions of the overall homework experience. As such the HRS-II may prove useful in the development of a model of the process of the client change process. The HRS-II provides a comprehensive assessment of the overall client experience of homework, covering the cognitive and behavioural factors found to be associated with homework completion.

In general, the results presented in Chapters 8 and 12 support the use of the HAACS and HRS-II in future research. Correlational anlyses with other measures suggest that the HRS-II and HAACS indices measure similar constructs tapped by other homework measures currently in use. Further, these results also indicated that the HAACS and HRS-II provide information not currently assessed by the comparison measures. For these reasons, it is expected that both the HAACS and HRS-II will provide a significant contribution to the Cognitive Behavior Therapy Homework Project aim of developing an understanding of the mechanism by which homework produces its effect in CBT.

References

Addis, M. E., & Jacobson, N. S. (2000). A closer look at the treatment rationale and homework compliance in cognitive-behavioral therapy for depression. *Cognitive Therapy and Research, 24*, 313-326.

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

Ajzen, I. (1991). The theory of planned behavior. *Organizational Behavior and Human Decision Processes, 50,* 179-211.

Alford, B. A., & Beck, A. T. (1994). Cognitive therapy of delusional beliefs. *Behaviour Research and Therapy, 32*, 313-326.

American Psychological Association (2001). *Publication manual of the American Psychological Association* (5th ed.). Washington: American Psychological Association.

Armitage, J. A. (2005). Can the theory of planned behaviour predict maintenance of physical activity? *Health Psychology, 24*, 235-245.

Ayres, J. J. B. (1998). Fear conditioning and avoidance. In W. O'Donohuue (Ed.), *Learning and behavior therapy*. Boston: Allyn & Bacon.

Bandura, A. (1977). Self-efficacy: Toward a unifying theory of behavioral change. *Psychological Review, 19*, 721 - 737.

Bandura, A. (1986). *Social foundations of thought and action: A social cognitive theory.* Englewood Cliffs, NJ: Prentice Hall.

Bandura, A. (1989). Human agency in social cognitive theory. *American Psychologist, 44*,

1175 -1184.

Barber, J. P., Crits-Christoph, P., & Luborsky, L. (1996). Effects of therapist adherence and competence on patient outcome in brief dynamic therapy. *Journal of Consulting and Clinical Psychology, 64,* 619–622.

Barber, J. P., Liese, B. S., & Abrams, M. J., (2003). Development of the cognitive therapy adherence and competence scale. *Psychotherapy Research, 6,* 12-24.

Barlow, D. H., & Durand, V. M. (2002). *Abnormal Psychology* (3rd ed). Belmont, CA:Wadsworth.

Beck A. T. (1983). Cognitive therapy of depression: New perspectives. In P. J. Clayton & J. E. Barrett (Eds.). *Treatment of depression: Old controversies and new approaches.* New York: Raven Press.

Beck, A. T. (1964). Thinking and depression: Theory and treatment. *Archives of General Psychiatry, 10,* 561-571.

Beck, A. T. (1976). *Cognitive therapy and the emotional disorders.* New York: International Universities Press.

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

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

Beck, J. S. (1994). Cognitive therapy: Basics and beyond. New York: Guildford Press.

Beck, J. S. (1995). *Cognitive therapy: basics and beyond.* New York: The Guilford Press.

Berk, L. E. (2001). *Development through the lifespan* (2nd ed.). Boston, MA: Allyn and Bacon.

Bjornholdt, A. (2006) *A preliminary psychometric investigation of the Homework Rating Scale-II.* Unpublished master's thesis, Massey University, Albany, New Zealand.

Bland, J. M., & Altman, D. G. (1986). Statistical methods for assessing agreement between two methods of clinical measurement. *The Lancet, 1*, 307-310.

Bland, J. M., & Altman, D. G. (2003). Applying the right statistics: Analyses of measurement studies. *Ultrasound, Obstetrics and Gynecology, 22*, 85-93.

Brown, G. P., Hammen, C. L., Craske, M. G., & Wickens, T. D. (1995). Dimensions of dysfunctional attitudes as vulnerabilities to depressive symptoms. *Journal of Abnormal Psychology, 3*, 431-435.

Bryant, M. J., Simons, A. D., & Thase, M. E. (1999). Therapist skill and patient variables in homework compliance: Controlling an uncontrolled variable in cognitive therapy outcome research. *Cognitive Therapy and Research, 23*, 381-399.

Bryk, A., & Raudenbush, S. W. (1992). *Hierarchical linear models for social and behavioral research: Applications and data analysis methods.* Newbury Park, CA: Sage.

Bryman, A., & Cramer, D. (2005). *Quantitative data analysis with SPSS 12 and 13: A guide for social scientists.* New York: Routledge.

Burns, D. D., & Nolen-Hoeksema, S. (1991). Coping styles, homework compliance, and the effectiveness of cognitive behavioral therapy. *Journal of Consulting and Clinical Psychology, 59*, 305 - 311.

Burns, D. D., & Spangler, D. (2000) Does psychotherapy homework lead to changes in depression in cognitive behavioral therapy? Or does clinical improvement lead to homework compliance? *Journal of Consulting and Clinical Psychology, 68,* 45-56.

Castonguay, L. J., & Holtforth, M. G. (2005). Change in psychotherapy: A plea for no more "nonspecific" and false dichotomies. *Clinical Psychology: Science and Practice, 12*, 198-201.

Cautela, J. R., & Kearney, A. (1993). *The covert conditioning handbook.* Pacific Grove, CA: Brooks Cole.

Chambless, D. L., Sanderson, W. C., Shoham, V., Johnston, S. B., Pope, K. S., Crits-Christoph, P., et al. (1996). An update on empirically validated treatments. *The Clinical Psychologist, 49*, 5- 18.

Chatoor, I., & Krupnick, J. (2001). The role of non-specific factors in treatment outcome of psychotherapy studies. *European Child & Adolescent Psychiatry, 10* (1), 19-25.

Cichetti, D, V., & Feinstein, A. R. (1990). High agreement but low kappa: II. Resolving the paradoxes. *Journal of Clinical Epidemiology, 43*, 543-549.

Cichetti, D. V. (1994) Guidelines, criteria, and rules of thumb for evaluating normed and standardized assessment instruments in psychology. *Psychological Assessment, 6*, 284-290.

Cichetti, D. V., & Sparrow, S. S. (1981). Developing criteria for establishing interrater reliability of specific items: Applications to assessment of adaptive behavior. *American Journal of Mental Deficiency, 84*, 127-137

Clark, D. A., Antony, M. M., Beck, A. T., Swinson, R. P., & Steer , R. A. (2005). Screening for obsessive and compulsive symptoms: Validation of the Clark-Beck Obsessive Inventory. *Psychological Assessment, 17*, 132-143.

Cliff, N. (1959). Adverbs as multipliers. *Psychological Review, 66*, 27-44.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ:

Erlbaim.

Cohen, J. (1992). A power primer. *Psychological Bulletin, 1*, 155-159.

Conoley, C. W., Padula, M. A., Payton, D. S., & Daniels, J. A. (1994). Predictors of client implementation of counselor recommendation: Match with problems, difficulty level and building on client strengths. *Journal of Counseling Psychology, 41*, 3 - 7.

Crits-Christoph, P., & Mintz, J. (1991). Implications of therapist effects for the design and analysis of comparative studies of psychotherapies. *Journal of Consulting and Clinical Psychology, 59*, 20-26.

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

Crits-Christoph, P., Gallop,R., (2006). Therapist effects in the National Institute of Mental Health Treatment of Depression Collaborative Research Program and other psychotherapy studies. *Psychotherapy Research, 16*, 178 –181.

Cronbach, L. J. (1971). Test validation. In R. L. Thorndike (Ed.), *Educational measurement* (2nd ed.). Washington, DC: American Council on Education.

Dancey, C. P., Reidy, J. (1998). *Statistics without maths: Using SPSS for Windows.*

Datillio, F. M. (2002) Homework assignments in couple and family therapy. *Journal of Clinical Psychology, 58*(5), 535 - 547.

Dattilio, F. M. (2002). Homework assignments in couple and family therapy. *Journal of Clinical Psychology, 58*(5), 535 - 547.

Davidson, K. M., Obonsawin, M. C., Seils, & Patience, L.,(2002). Patient and

clinician agreement using the SWAP-200. *Journal of Personality Disorders, 17*, 208-218.

DeLeon, I. G., & Iwata, B. A. (1996). Evaluation of a multiple-stimulus presentation format forassessing reinforcer preferences. *Journal of Applied Behavioural Analysis, 25*, 519- 533.

Detweiler, J. B., & Whisman, M. A. (1999). The role of homework assignments in cognitive therapy for depression: Potential methods for advancing adherence. *Clinical Psychology: Science and Practice, 6*, 267-282.

Detweiler-Bedell, J. B., & Whisman, M. A. (2005). A lesson in assigning homework: Therapist, client, and task characteristics in cognitive therapy for depression. *Professional Psychology: Research and Practice, 36*(2), 319 - 223.

Dewitte, K., Fierens, C., Stockl, D., & Thienpont, L. M. (2002). Application of the Bland-Altman plot for interpretation of method-comparison studies: A critical investigation of its practice. *Chemistry, 48*, 799-801.

Dimidjian, S., Dobson, K. S., Kohlenberg, R. J., Gallop, R., Markley, D. K., Atkins, D. C., et al. (in press). Randomized trial of behavioral activation, cognitive therapy, and antidepressant medication in the acute treatment of adults with major depression. *Journal of Consulting and Clinical Psychology.*

Dobson, J. S. Beck, & A. T. Beck (2005). Academy of cognitive therapy: Purpose, history, and future prospects. *Cognitive and Behavioral Practice, 12*, 263-266.

Dobson, K. S., & Kazantzis, N., (2003). The therapist in cognitive-behavioral therapy: Introduction to a special section. *Psychotherapy Research, 13*, 131-134.

Dobson, K. S., & Singer, A. R. (2005). Definitional and practical issues in the assessment of treatment integrity. *Clinical Psychology: Science and practice, 12,* 384-387.

Doss, B, D. (2004). Changing the Way We Study Change in Psychotherapy. *Clinical Psychology: Science and Practice, 11*, 368-386.

Edelman, R. E., & Chambless, D. L. (1995). Adherence during sessions and homework in cognitive-behavioral group treatment of social phobia. *Behaviour Research and Therapy, 33*, 573 - 577.

Elkin, I. Shea, M. T., Watkins, J. T., Imber, S. D., Sotsky, S. M., Collins, J. F., et al. (1989). National Institute of Mental Health Treatment of Depression Collaborative Research Program: General effectiveness of treatments. *Archives of General Psychiatry, 46,* 971-982.

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

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

Elliott, R., Watson, J.C., Greenberg, L., & Goldman, R. (2003).Learning emotion focused psychotherapy: The process-experiential approach to change. Washington: APA Books.

Ellis, A., & Dryden, W. (1997). *The practice of rational emotive behavior therapy*. New York: Springer Publishing.

Ellis, A., & Dryden, W. (1997). The practice of rational emotive behavior therapy. New York: Springer Publishing.

Feeley, M., DeRubeis, R. J., & Gelfand, L. A. (1999). The temporal relation of adherence and alliance to symptom change in cognitive therapy for depression. *Journal of Consulting and Clinical Psychology, 67*, 578-582.

Feinstein, A. R., & , D. V. (1990). High agreement but low kapa: I. The problem of two paradoxes. *Journal of Clinical Epidemiology, 43*, 543-549.

Feldt, L. S., & Brennan, R. L. (1989). Reliability. In R. L. Linn (Ed.). *Educational measurement* (3rd ed.). Phoenix, AZ: Oryx Press.

Fennell, M. J. V., & Teasdale, J. D. (1987). Cognitive therapy for depression: Individual differences and the process of change. *Cognitive Therapy and Research, 11*, 253 - 271.

Festinger, L. (1957). *A theory of cognitive dissonance.* Stanford: Stanford University Press.

Fleiss, J. L. (1981). *Statistical methods for rates and proportions* (2nd ed.) New York: Wiley.

Floyd, D. L., Prentice-Dunn, S., & Rogers, R. W. (2000). Ameta-analysis of research on protection motivation theory. *Journal of Applied Social Psychology, 30*, 407–429.

Foley, S. H., O'Malley, S., Rounsaville, B. J., Prusoff, B. A., & Weissman, M. M. (1987). The relationship between patient difficulty and therapist performance in interpersonal psychotherapy of depression. *Journal of Affective Disorders, 12*, 207-217.

Friedberg, R. D., & McClure, J. M. (2005). Adolescents. In N. Kazantzis, F. D. Deane, K. R. Ronan & L. L'Abate (Eds.), *Using homework assignments in cognitive-behaviour therapy* (pp. 95 -116). New York: Brunner-Routledge.

Gilboa-Schechtman, E., & Shahar, G. (2006). The sooner the better: Temporal patterns in brief treatment of depression and their role in long-term outcome.

Psychotherapy Research, 16, 374-384.

Gollwitzer, P. M. (1993). Goal achievement: The role of intentions. *European Review of Social Psychology, 4*, 141–185.

Green,.S. B., & Salkind, N. J. (2005). *Using SPSS for windows and macintosh: Analizing and understanding data* (4th ed.) Upper Saddle River, NJ: Pearson Prentice Hall.

Greenberg, L. S., Watson, J. C., & Goldman, R. (1988). *Process-experiential therapy of depression: Handbook of experiential psychotherapy*. New York: Guilford Press.

Gresham, F. M. (2005). Treatment integrity and therapeutic change: Commentary on Perepletchikiva and Kazdin. *Clinical Psychology: Science and Practice*, 391-394.

Haggard, E. A. (1958). *Intraclass correlation and the analysis of variance*, New York: Dryden.

Hair,J. F., Anderson, R. E., Tatham, R. L., & Black, W. C. (1998). *Multivariate data analysis* (5th ed.). Upper Saddle River, NJ: Prentice Hall.

Hall, S. F., Groome, P. A., Streiner, D. L., Rochon, P. A. (in press). Interrater reliability and measurements of comorbid illness. *Journal of Clinical Epidemiology*.

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

Hawkins, D. M. (2002). Diagnostics for conformity of paired quantitative measurements. *Statistics in Medicine, 21*, 1913-1935.

Haynes, S. N., Richard, C. S., & Kubany, E. S. (1995). Content validity in psychological assessment: A functional approach to concepts and methods. *Psychological Assessment, 7*, 238-247.

Heckhausen, H. (1991). *Motivation and action*. New York: Springer.

Helbig, S., & Fehm, L. (2004). Problems with homework in CBT: Rare exception or rather frequent? *Behavioural and Cognitive Psychotherapy, 32*, 291-301.

Herzberg, A. (1941). Short-term treatment of neurosis by graduate tasks. *British journal of medical psychology, 29*, 36 - 51.

Hill, C. E., Lambert, M. J. Methodological issues in studying psychotherapy processes and outcomes. In M. J. Lambert (Ed.). *Bergin and Garfield's Handbook of psychotherapy and behavior change* (fth ed.). New York: John Wiley & Sons.

Hill, C. E., O'Grady, K. E., & Elkin, I. (1992), Applying the collaborative study psychotherapy rating scale to rate therapist adherence to cognitive-behavior therapy, interpersonal therapy, and clinical management. *Journal of Consulting and Clinical Psychology, 50*,73-79.

Hoberman, H. M., & Lewisohn, P. M. (1985). The behavioral treatment of depression. In E. E. Beckham & W. R. Lember (Eds.), Handbook of depression: Treatment, assessment, and research (pp. 39-91). Homewood, IL: Dorsey Press.

Hogan, T. P, & Agnello, J. (2004). An empirical study of reporting practices concerning measurement validity. *Educational and Psychological Measurement, 64*, 802-812.

Hong, E., & Milgram, R. A. (2000). *Homework: Motivation and learning preference.* Westport, CT: Bergin & Garvey.

Howe, E. S. (1966). Associative structure of quantifiers. *Journal of Verbal Learning and Verbal Behavior, 5*, 156-162.

Hunsley, J., Dobson, K. S., Johnston, C., & Mikail, S. F. (1999). Empirically supported treatments in psychology: Implications for Canadian professional psychology. *Canadian Psychology, 40*, 289-302.

Ilardi, S. S., & Craighead, W. E. (1994). The role of nonspecific factors in cognitive-behavior therapy for depression. *Clinical Psychology: Science and Practice, 1*, 138-156.

Imber, S.D., Pilkonis, P. A.,Sotsky, S. M., Elkin, I., Watkins, J. T., et al. (1990). Mode-specific effects among three treatments for depression. *Journal of Consulting and Clinical Psychology, 58*, 352-359.

Jacobson, N. S. (1998, July). *How important are cognitive interventions in cognitive therapy?* Paper presented at the World Congress of Behavioral and Cognitive Therapies, Acapulco, Mexico.

Jacobson, N. S., Dobson, K. S., Truax, P. A., Addis, M. E., Koerner, K., Gollon, J. K., et al. (1996). A component analysis of cognitive-behavioral treatment for depression. *Journal of Consulting and Clinical Psychology, 64*, 295-304.

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

Johnson, M.H., Kazantzis, N. (2004) Cognitive behavioral therapy for chronic pain: strategies for the successful use of homework assignments. Journal of Rational Emotive and Cognitive- Behavioural Therapy 22, 187-216.

Kaiser, H. (1970). A second generation little jiffy. *Psycometrika, 35*, 401-415.

Kaiser, H. (1974). An index of factorial simplicity. *Psychometrika, 39*, 31 - 36.

Kanfer, F. H., & Phillips, J. S. (1966). A survey of current behavior and a proposal for classification. *Archives of General Psychiatry, 15*, 114-128.

Kaplan, R. M., & Saccuzzo (2001). *Psychological testing: Principles, applications, and issues* (5th ed.). Belmont, CA: Wadswoth/Thompson.

Kazantzis, N. (2000). Power to Detect Homework Effects in Psychotherapy Outcome Research.*Journal of Consulting and Clinical Psychology, 68*(1), 166-170.

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

Kazantzis, N., & Deane, F. D. (1999). Psychologist's use of homework assignments in clinical practice. *Professional Psychology: Research and Practice, 30*, 581 - 585.

Kazantzis, N., & L'Abate, L. (2005). Theoretical foundations. In N. Kazantzis, F. D. Deane, K. R. Ronan & L. L'Abate (Eds.), *Using homework assignments in cognitive-behaviour therapy* (pp. 9 - 34). New York: Brunner-Routledge.

Kazantzis, N., & Lampropoulos, G. K. (2002). Reflecting on homework in psychotherapy: What can we conclude from research and experience? *Journal of Clinical Psychology, 58*(5), 577 - 584.

Kazantzis, N., Busch, R., Ronan, K. R., & Merrick, P. L. (2004). Mental health practitioners' use and perceived importance of homework assignments in psychotherapy. *Journal of Consultingand Clinical Psychology, 68*, 166 - 170.

Kazantzis, N., Dattilio, F. M., & Merrick, P. L. (in press). Do Researchers Do Their Homework? A 10-Year Survey of Psychotherapy Outcome Research Published in Journal of Consulting and Clinical Psychology. *Journal of Consulting and Clinical Psychology*.

Kazantzis, N., Deane, F. D., & Ronan, K. R. (2000). Homework assignments in cognitive and behavioral therapy: A meta-analysis. *Clinical Psychology: Science and Practice, 7*, 189 - 202.

Kazantzis, N., Deane, F. D., & Ronan, K. R. (2004). Assessing compliance with homework assignments: Review and recommendations for clinical practice.

Journal of Clinical Psychology, 60(627 - 641).

Kazantzis, N., Deane, F. D., & Ronan, K. R. (2005). Assessment of homework completion. In N. Kazantzis, F. D. Deane, K. R. Ronan & L. L'Abate (Eds.), *Using homework assignments in cognitive-behaviour therapy* (pp. 61 - 74). New York: Brunner-Routledge.

Kazantzis, N., Deane, F. D., Ronan, K. R., & Lampropoulos, G. K. (2005). Empirical foundations. In N. Kazantzis, F. D. Deane, K. R. Ronan & L. L'Abate (Eds.), *Using homework assignments in cognitive-behaviour therapy* (pp. 35 - 60). New York: Brunner-Routledge.

Kazantzis, N., Deane, F. P., & Ronan, K. R. (2005). *Homework Rating Scale II (HRS-II).*

Kazantzis, N., Deane, F. P., Lamploproulos, G. K. (2005). A national survey of practicing psychologists' use and attitudes toward homework in psychotherapy. *Journal of Consulting and Clinical Psychology, 73,* 742-748.

Kazantzis, N., Lampropoulos, G. K., & Deane, F. P. (2005). A national survey of practicing psychologists' use and attitudes toward homework in psychotherapy. *Journal of Consulting and Clinical Psychology, 73*(4), 742 - 748.

Kazantzis, N., MacEwen, J., & Datillio, F. (2005). A guiding model for practice. In N. Kazantzis, F.P. Deane, K. R. Ronan (Eds.), *Using homework Assignments in Cognitive-Behavior Therapy*. New York: Routledge.

Kazantzis, N., & Ronan, K. R. (2006). Can between-session (homework) activities be considered a common factor in psychotherapy? *Journal of Psychotherapy Integration, 16,* 115-127.

Kazantzis, N., Wedge, P., & Dobson, K. S. (2005). *Homework Adherence and Competence Scale (HAACS).*

Kazdin, A. E. (2003). Special topics in clinical research: Introduction. In A. E. Kazdin (Ed.), *Methodological issues and strategies in clinical research* (3rd ed.) (493-501). Washington, D. C.: American Psychological Association.

Kazdin, A. E., & Mascitelli, S. (1982). Covert and overt rehearsal and homework practice in developing assertiveness. *Journal of Consulting and Clinical Psychology, 50*, 250 - 258.

Kendall, P. C., Marrs-Garcia, A., Nath, S. R., & Sheldrick, R. C. (1999). Normative comparisons for the evaluation of clinical significance. *Journal of Consulting and Clinical Psychology, 67*, 285-299.

Kim J-O., & Mueller C.W. (1978). *Introduction to factor analysis: What it is and how to do it.* Beverly Hills, CA: Sage Publications.

Kline, P. (1998). *The new psychometrics: Science, psychology, and measurement.* London: Routledge.

Kohler, F. W., & Greenwood, C. R. (1986). Toward technology of generalization: The identification of natural contingencies of reinforcement. *The Behaviour Analyst, 9*, 19 - 26.

Lambert, M. J., Harmon, S. C., & Slade, K. (in press). Directions for research on homework. In N. Kazantzis & L. L'Abate (Eds.), *Handbook of homework assignments in psychotherapy: Research, practice and prevention* (pp. 709 - 742). New York: Springer.

Lampropoulos, G. K., & Rector, N. A. (2004). *Homework compliance in cognitive behavior therapy for anxiety disorders: Predictors of outcome.* Paper presented at the Poster presented at the 30th Annual Harvey Stancer Research Day, Department of Psychiatry, University of Toronto, Toronto.

Landis J, & Koch G. (1977), The measurement of observer agreement for categorical data. *Biometrics.33*, 159-174.

Lau, M. A., Dubord, G. M., Parikh, S. V. (2004). Design and feasibility of a new cognitive-behavioural therapy course using a longitudinal interactice format. *Canadian Journal of Psychiatry, 49*, 696-700.

Leahy, R. L. (2001). *Overcoming resistance in Cognitive Therapy*. New York: Guilford Press.

Leahy, R. L. (2002). Improving homework compliance in the treatment of generalized anxiety disorder. *Psychotherapy in Practice, 58*, 499-511.

Looney, M. A. (2000). When is the intraclass correlation coefficient misleading? *Measurement in Physical Education and Health Science, 4*, 73-78.

MacCallum, R. C., Widaman, K. F., Zhang, S., & Hong, S. (1999). Sample size in factor analysis. *Psychological Methods, 4*, 84-99.

Mahrer, A. R., Nordin, S., & Miller, L. S. (1995). If a client has this kind of problem, prescribe that kind of post-session behavior. *Psychotherapy, 32*, 194 - 203.

Martin, G., & Pear, J. (2003). *Behavior modification: What it is and how to do it* (7th ed.). Upper Saddle River, NJ: Prentice Hall.

Masters, H., & Johnson, E. (1970). *Human sexual inadequacy*. Boston: Little Brown & Company.

McGlinchey, J. B., & Dobson, K. S. (2003). Treatment integrity concerns in Cognitive Therapy for depression. *Journal of Cognitive Psychotherapy, 17*, 299-318.

McIntosh, V. W., Jordan, J., McKenzie, J. M., Luty, S. E., Carter, F. A., Carter, J. D., et al. (2005). Measuring therapist adherence in psychotherapy for anorexia nervosa: Scale adaptation, psychometric properties, and distinguishing psychotherapies. *Psychotherapy Research, 15,* 339-344.

McLeod, B. & Weisz, J. R. (2005). The therapy process observational coding system-Alliance scale: Measure characteristics and prediction of outcome in usual clinical practice. *Journal of Consulting and Clinical Psychology, 73*, 323-333.

Meier S. T. & Davis S. R. (1990). Trends in reporting psychometric properties of scales used in counseling psychology research. *Journal of Counseling Research, 37*, 113-115.

Messick, S. (1993). Validity. In R. L. Linn (Ed.), *Educational measurement* (3rd ed.). Phoenix, AZ: Oryx Press.

Messick, S. (1995). Validity of psychological testing: Validation of inferences from persons' responses and performances as scientifiv inquiry into score meaning. *American Psychol;ogist, 50*, 741-749.

Mischel W. (1973). Toward a cognitive social learning reconceptualization of personality. *Psychological Review, 80*, 252–283

Miller, S. J., & Binder, J. L. (2002). The effects of manual-based training on treatment fidelity and outcome: A review of the literature on adult individual psychotherapy. *Psychotherapy: Theory, Research, Practice, and Training, 39*, 184–198.

Milne, D. L. & James , I. A. (2002). The observed impact of training on competence in clinical supervision. *British Journal of Clinical Psychology, 41*, 55-72.

Milne, D. L. (Ed.) (1987). *Evaluating mental health practice: Methods and applications.* London: Croom-Helm.

Milne, D., Woodward, K., Hanner, S., Iceton, J., Fitzsimmons, A., & Rochester, J. (2003). An illustration of delivering evidence-based practice through staff training: Multi-dimensional process, outcome and organizational evaluation. *Behavioural and Cognitive Psychotherapy, 31*, 85-98.

Milne, S., Orbell, S., & Sheeran, P., (2002) Combining motivational and volitional interventions to promote exercise participation: Protection motivation theory and implementation intentions. *British Journal of health psychology, 7,* 163-184.

Moncher, F. J., & Prinz, R. J. (1991). Treatment fidelity in outcome studies. *Clinical Psychology Review, 11*, 247-266.

Needleman, L. D., (1999). *Cognitive case conceptualization: A guidebook for practitioners.* Mahwah, NJ: Lawrence Erlbaum Associates.

Neimeyer, R. A., & Feixas, G. (1990). The role of homework and skill acquisition in the outcome of group cognitive therapy for depression. *Behaviour Therapy, 21*, 281 - 292.

Nezu, A. M., & Nezu, C. M. (2005). Comments on "Evidence-based behavioral medicine: What is it and how do we achieve it?": The interventionist does not always equal the intervention – The role of therapist competence. *Annals of Behavioral Medicine, 29*(1), pp. 80.

Norusis, M. J. (2006). *SPSS 13.0 statistical procedures companion.* Upper Saddle River: Prentice Hall.

Nunnally, J. C., Bernstein, I. H. (1994). *Psychometric theory* (3rd ed.). New York: McGraw-Hill.

Okiishi, J., Lambert, M. J., Nielsen, S. L., & Ogles, B. M. (2003). Waiting for supershrink: An empirical analysis of therapist effects. *Clinical Psychology and Psychotherapy, 10*, 361-373.

Pachana N. A., & Sofronoff, K. (2005). Sexual problems. In N. Kazantzis, F. D. Deane, K. R. Ronan & L. L'Abate (Eds.), *Using homework assignments in cognitive-behaviour therapy* (pp. 307-328). New York: Brunner-Routledge.

Padesky, C. A., & Greenberger, D. (1995). *Clinician's guide to Mind over Mood.* New York: Guilford Press.

Park, J. M., Mataix-Cols, D., Marks, I. M., Ngamthipwatthana, T., Marks, M., Araya, R., et al. (2001). Two-year follow-up after a randomized controlled trial of self- and clinician-accompanied exposure for phobia/panic disorders. *British Journal of Psychiatry, 178*, 543 -548.

Perepletchikova, & Kazdon, A. E. (2005). Treatment integrity and therapeutic change: Issues and research recommendations. *Clinical Psychology: Science and Practice, 12,* 365-383.

Persons, J. B. (1989). *Cognitive therapy in practice: A case formulation approach.* New York: W. W. Norton & Company.

Persons, J. B., & Bertagnolli, A. (1999). Inter-rater reliability of Cognitive-Behavioral case formulations of depression: A replication. *Cognitive Therapy and Research, 23*, 271-283.

Persons, J. B., Davidson, J., & Tompkins, M. A. (2001). *Essential componenets of cognititve-behavior therapy for depression.* Washington: American Psychological Association.

Pretorius, W. M. (2006), Cognitive Behavioural Therapy Supervision: Recommended Practice. Behavioural and Cognitive Psychotherapy. Retrieved from http://journals.cambridge.org. doi:10.1017/S1352465806002876.

Primakoff, L., Epstein, N., & Covi, L. (1986). Homework compliance: An uncontrolled variable in cognitive therapy outcome research. *Behavior Therapy, 17*, 433- 446.

Primakoff, L., Epstein, N., & Covi, L. (1986). Homework compliance: An uncontrolled variable in cognitive therapy outcome research. *Behaviour Therapy, 17*, 433-446.

Prochaska, J. O. (1979). Symptom and dynamic cues in the implosive treatment of test anxiety. Journal of Abnormal Psychology, 77, 133-142.

Prochaska, J. O., DiClemente, C. C., & Norcross, J. C. (1992). In search of how people change: Applications to addictive behaviors. *American Psychologist, 47*(9), 1102 - 1114.

Prochaska, J. O., DiClemente, C. C., Norcross, J. C. (1992). In search of how people change. *American Psychologist, 47*, 1102-1114.

Prochaska, J. O., Velicer, W. F., Rossi, J. S., Goldstein, M. G., Marcus, B. H., Rakowski, W., et al. (1994). Stages of change and decisional balance for 12 problem behaviors. *Health Psychology, 13*, 39 - 46.

Qualls A. L., & Moss, A. D. (1996). The degree of congruence between tes standards and test documentation within journal publications. *Educational an Psychological Measurement, 56*, 209-214.

Rachman, S. J. (1990). *Fear and courage*. New York: Freeman.

Rankin, G., & Stokes, M. (1998). Reliability of assessment tools in rehabilitation: An illustration of appropriate statistical analyses. *Clinical Rehabilitation, 12,* 187-199.

Reicheldt, F. K., James, I. A., & Blackburn, I-M (2003). Impact of training on rating competence in cognitive therapy.

Rogers, R. W. (1983). Cognitive and physiological processes in fear appeals and attitude change: A revised theory of protection motivation. In J. T. Cacioppo & R. E. Petty (Eds.), *Social Psychology* (pp. 153-176). New York: Guilford.

Rogers, R. W.. (1975).A protection motivation theory of fear appeals and attitude change. *Journal of Psychology, 91*, 93-114.

Savin N. E., & White, K. J. (1977). The Durbin-Warson test for serial correlation with extreme sample sizes or many regressors. *Econometrika, 45*, 1899-1996.

Scheel, M. J., Hanson, W. E., & Razzhavaikina, T. I. (2004). The process of recommending homework in psychotherapy: A review of therapist delivery methods, client acceptability, and factors that affect compliance. *Psychotherapy: Theory, Research, Practice, Training, 41*, 38-55.

Schmidt, N. B., & Woolaway-Bickel, K. (2000). The effects of treatment compliance on outcome in cognitive-behavioral therapy for panic disorder: Quality versus quantity. *Journal of Consulting and Clinical Psychology, 68*, 13 - 18.

Schmidt, N. B., & Woolaway-Bickel, K. (2000). The effects of treatment compliance on outcome in cognitive-behavioral therapy for panic disorder: Quality versus quantity. *Journal ofConsulting and Clinical Psychology, 68*, 13 - 18.

Shaw, B. F., Elkin, I., Yamaguchi, J., Olmsted, M., Vallis, T. M., Dobson, K. S., et al. (1999). Therapist competence ratings in relation to clinical outcome in cognitive therapy of depression. *Journal of Consulting and Clinical Psychology, 67*, 837–846.

Shaw, B. F., Elkin, I., Yamaguchi, J., Olmsted, M., Vallis, T. M., Dobson, K. S., et al. (1999). Therapist Competence Ratings in Relation to Clinical Outcome in Cognitive Therapy of Depression. *Journal of Consulting and Clinical Psychology, 67*, 837-846.

Sheeran, P.,Trafimow, D.,& Armitage, C. J. (2003). Predicting behaviour from perceived behavioural control: Tests of the accuracy assumption of the theory of planned behaviour. *British Journal of Social Psychology, 42*, 393-410.

Shelton J. L., & Ackerman, J. M. (1974). *Homework in counseling and psychotherapy: Examples of systematic assignments for therapeutic use by mental health professionals.* Springfield, IL: Charles C. Thomas.

Shelton, J. L., & Levy, R. L. (1981b). *Behavioral assignments and treatment compliance: A handbook of clinical strategies.* Champaign, IL: Research Press.

Shrout, P. E., & Fleiss, J. L. (1979). Intraclass correlations: Uses in assessing rater reliability. *Psychological Bulletin, 86*, 420-428.

Simons, A. D., Murphy, G. E., Levine, J. L., & Wetzel, R. D. (1986). Cognitive therapy and pharmacotherapy for depression: Sustained improvement over one year. *Archives of General Psychotherapy, 25*, 37 - 45.

Soldze, S., & McCullough, L. (2000). Introduction. In S. Soldz, & L. McCullough (Eds.), *Reconciling empirical knowledge and clinical experience* (pp. 3-13). Washington, D.C.: American Psychological Association.

Springer, C., Reddy, L. (2004). Measuring adherence in behavior therapy: Opportunities for practice and research. *The Behavior Therapist, 27*, 99-100.

SPSS. (2004). Statistical Package for the Social Sciences (Version 13): SPSS, inc.

Sreiner, D. L. (2003a). Being inconsistent about consistency: When coefficient alpha does and doesn't matter. *Journal of Personality Assessment, 80*, 217-222.

Sreiner, D. L. (2003b). Starting at the beginning: An introduction to coefficient alpha and internal consistency. *Journal of Personality Assessment, 80*, 99-103.

Startup, M., & Edmonds, J. (1994). Compliance with homework assignments in

cognitive-behavioral psychotherapy for depression: Relation to outcome and methods of enhancement. *Cognitive Therapy and Research, 18*, 567 - 579.

Sternberg, R. J. (1996). *Cognitive psychology*. Orlando, FL: Harcourt Brace College Publishers.

Stokes, T. F., & Baer, D. M. (1977). An implicit technology of generalization. Journal of Applied Behavior Analysis, 10, 349-367.

Streiner, D. L. (2003a). Being Inconsistent About Consistency: When Coefficient Alpha Does and Doesn't Matter. *Journal of Personality Assessment, 80*, 217–222

Streiner, D. L. (2003b). Starting at the beginning: An introduction to coefficient Alpha and internal consistency. *Journal of Personality Assessment, 80*, 99–103.

Streiner, D. L., (1993). An introduction to multivariate statistics. *Canadian Journal of Psychiatry, 38*, 9-13.

Streiner, D. L., (2002). The case of the missing data: Methods of dealing with dropouts and other research vagaries. Canadian Journal of Psychiatry, 47, 68-75.

Sudak, D., Beck, J. S., & Wright, J. (2003). Cognitive behavioral therapy: A blueprint for attaining and assessing psychiatry resident competence. *Academic Psychiatry, 27*, 154-159.

Suen, H. K. (1988). Agreement, reliability, accuracy, and validity: Towards a clarification. *Behavioral Assessment, 10*, 343-366.

Svartberg, M. (1999). Therapist competence: Its temporal course, temporal stability, and determinants in short-term anxiety-provoking psychotherapy. *Journal of Clinical Psychology, 55*, 1313-1319.

Tang, T. Z., & DeRubeis, R. J. (1999). Sudden gains and critical sessions in cognitive-behavior therapy for depression. *Journal of Consulting and Clinical Psychology, 67*, 894-904.

Tang, T. Z., Beberman, R., DeRubeis, R. J. & Pham, T. (2005). Cognitive change, critical sessions, and sudden gains in cognitive-behavioral therapy for depresson. *Journal of Consulting and Clinical Psychology, 73*, 168-172.

Taylor, S. (1995). Assessment of obsessions and compulsions: Reliability, validity, and sensitivity to treatment effects. Clinical Psychology Review, 15, 261-296.

Thase, M. E., & Callan, J. A. (2006). The role of homework in cognitive behavior therapy of depression. *Journal of Psychotherapy Integration, 16*, 162-177.

Thompson, B. (2003) Understanding reliability and coefficient alpha, really. In B. Thompson (ed.). *Score reliability: Contemporary thinking on reliability issues.* Thousand Oaks, CA: Sage Publications.

Thompson, B. (2004). *Exploratory and confirmatory factor analysis.* Washington: Amereican Psychological Association.

Thorndike, E. L. (1927). The law of effect. *American journal of psychology, 39*, 212 - 222.

Tompkins, M. A. (2002). Guidelines for enhancing homework. *Journal of Clinial Psychology, 58*, 565- 576.

Tompkins, M. A. (2003). Effective homework. In R. L. Leahy (Ed.). *Roadbnlocks in cognitive behavior therapy: Transforming challenges into opportunities for change* (pp. 49-69). New York: Guilford Press.

Tompkins, M. A. (2004). *Using homework in psychotherapy: Strategies, guidelines, and forms.* New York: Guilford Press.

Uebersax, J. S. (1987). Diversity of decision-making models and the measurement of interrater agreement. P*sychological Bulletin, 101*, 140-146.

Vallis, T. M., Shaw, B. F., & Dobson, K. S. (1986). The Cognitive Therapy Scale: Psychometric properties. *Journal of Consulting and Clinical Psychology, 54,* 381–385.

Vallis, T. M., Shaw, B. F., & Dobson, K. S. (1986, The Cognitive Therapy Scale: Psychometric Properties. *Journal of Consulting and Clinical Psychology, 54,*

Vermilyea, B. B., Barlow, D. H., & O'Brien, G. T. (1984). The importance of assessing treatment integrity and outcomes in wraparound services. Journal of Child and Family Studies, 9, 351-361.

Waltz, J., Addis, M. E., Koerner, K., & Jacobson, N. S. (1993). Testing the integrity of a psychotherapy protocol: Assessment of adherence and competence. *Journal of Consulting and Clinical Psychology, 61,*

Wampold, B. E., & Bhati, K. S. (2004). Attending to the omissions: A historical examination of evidence-based practice movements. *Professional Psychology: Research and Practice, 35*, 563-570.

Wedge, P. (2005), The development of the Homework Adherence and Competence scale (HAACS): A measure for assessing therapist adherence and competence in administering homework assignments within cognitive behaviour therapy. Unpublished master's thesis, Massey University, Albany, New Zealand.

Weerasekera, P., Antony, M. M., Bellisimo, A., Bieling, P., Shurina-Egan, J., Spenser, A., et al. (2003). Competency assessment in the McMaster psychotherapy program. *Academic Psychiatry, 2003, 27*, 166-173.

Whittal, M. L.,Thordarson, D. S., & McLean,P. D. (2005). Treatment of obsessive-

compulsive disorder: Cognitive behavior therapy vs. exposure and response prevention. *Behaviour Research and Therapy, 43*, 1559-1576.

Whittington, D., (1998). How well do researchers report their measures? *Educational and Psychological Measurement, 58*, 21-37.

Wilkinson, L., & American Psychological Association Board of Scientific Affairs Task Force on Statistical Inference. (1999). Statistical methods in psychology journals: Guidelines and explanations. *American Psychologist, 54*, 594 - 604.

Wood, D. (1998). *How children think and learn: The social contexrs of cognitive development* (2nd ed.). Oxford: Bkackwell.

Woods, D. (1998). How Children think and learn. Blackwell:Oxford.

Woody, S. R., & Adessky, R. S. (2002). Therapeutic alliance, group cohesion, and homework compliance during cognitive behavioral group treatment of social phobia. *Behavior Therapy, 33*, 5–27.

Wright, J. H., Thase, M. E., Beck, A. T., & Ludgate, J. W. (1993). *Cognitive therapy with inpatients: Developing a cognitive milieu.* New York: Guilford Press.

Young, J. E., Klosko, J. S., & Weishar, M. E. (2003). *Schema therapy: A practitioner's guide.* New York: Guilford.

Young, J., & Beck, A. T. (1980). Cognitive therapy scale: Rating manual, (Available from the Beck Institute for Cognitive Therapy and Research, GSB Building, 1 Belmont Avenue, Suite 700, Bala Cynwyd, PA 19004-1610). 38.academyofct.org/upload/documents/CTRS_Manual.pdf

APPENDIX A

HOMEWORK ADHERENCE AND COMPETENCE SCALE

HAACS

Homework Adherence and Competence Scale

Nikolaos Kazantzis
Paul Wedge
Keith S. Dobson

Second Draft

HAACS Homework Adherence and Competence Scale

Instructions:
This therapist adherence and competence rating scale consists of 19 items regarding therapists' integration of homework assignments in cognitive behavior therapy (CBT). Items 1-5 cover therapist behaviors in REVIEWING previously assigned homework. Items 6-14 cover therapist behaviors in DESIGNING new or revised homework. Items 15-19 cover therapist behaviors in ASSIGNING how the new or revised homework will be practically carried out. Please note that although the items are categorized into these three conceptually different groupings, they are often not so clearly delineated during a CBT session. Finally, each individual section (i.e., review, design and assign) concludes with an overall rating for that section.

Every individual item begins on a new page, and has two clearly identifiable questions; adherence (e.g., "DID the therapist ..." or "WAS a ...") and competence (i.e. "HOW WELL did the therapist ...") The adherence question for each item is labeled with an 'a' (e.g., 1a, 2a, etc.), and the competence question for each item is labeled with a 'b' (e.g., 1b, 2b, etc.).

(a) **Adherence**
Please note that your rating for the *adherence* questions (i.e., the 'a' questions) is to indicate whether these aspects were carried out in the session to any extent. This is different from rating how well the therapist undertook each item (i.e., competence). For each adherence item, please consider the question carefully, and tick either "yes" or "no" to indicate whether the particular aspect was CARRIED OUT irrespective of how well it was done. Please select only one response option for any question.

(b) **Competence**
Please note that your rating for the *competence* questions (i.e., the 'b' questions) is to indicate HOW WELL the therapist undertook each item. This is different from rating whether these aspects were carried out by the therapist (i.e., adherence). Adherence is a necessary BUT NOT SUFFICIENT condition for competence. This means that if adherence was rated "no" for a therapist behavior, then the therapist competence cannot be rated higher than "0" for the same item. Conversely, any competence rating between 1 and 6 necessitates a "yes" adherence rating. These "rules" provide a double check that you are rating adherence and competence correctly. For each item, please consider the competence question carefully and record the appropriate number in the rating box to indicate how well the therapist carried out each aspect.

Each competence question has seven descriptive response options. In the first instance, please use the descriptive response options to determine the rating for each item. Please be aware that these response options build in complexity from 0 to 6, with each increment adding more complex or additional requirements. To qualify for the higher rating, then all of the components of that descriptive response option must be met. If this is not the case then please revert to the next lowest option in which the criteria are fully met. However, if you are having difficulty deciding on a rating (e.g., the response options descriptions do not seem to easily fit the session being rated), then use the 7-point Likert scale ranging from 0 (*non-adherence/extremely poor*) to 6 (*excellent*).

0	1	2	3	4	5	6
Non-adherence/ extremely poor	**Poor**	**Mediocre**	**Fair**	**Good**	**Very Good**	**Excellent**

If several items seem to apply equally well, record the lowest number (e.g., if considering recording "3-4", record it as a "3"). Please provide a single rating for every item.

HOMEWORK REVIEW Items 1-5 cover the therapist behaviors involved in reviewing homework from the previous session, and typically occurs early in the session.

0	1	2	3	4	5	6
Non-adherence/ extremely poor	Poor	Mediocre	Fair	Good	Very Good	Excellent

Item 1

1a **DID the therapist discuss the completion of previously assigned homework to any extent?** Yes ☐ No ☐

1b **HOW WELL did the therapist discuss the completion of previously assigned homework?** Competence Rating ☐

0 The therapist DID NOT discuss previously assigned homework.

1 The therapist made a CURSORY ENQUIRY about previous homework completion, but DID NOT ENGAGE the client (i.e., no exploration of the client's responses).

2 The therapist ENQUIRED about previous homework completion, and made an attempt to elicit feedback from the client but this was NOT SUCCESSFUL (e.g., the therapist used closed questions, or did not allow sufficient time for a response).

3 The therapist ENQUIRED about previous homework, and elicited some GENERAL FEEDBACK from the client. For instance, the client gave a vague response such as "I completed most of it" and this response was taken at face value and was not explored further (e.g., "Can you tell me more about the parts you completed?" and then "Can you tell me about the parts you had difficulty with or did not complete?").

4 The therapist ENQUIRED about previous homework and IDENTIFIED EXACTLY what portion of the homework was completed and what was not completed. However, the discussion focused EITHER on the completed homework OR the non-completed homework.

5 The therapist IDENTIFIED and DISCUSSED BOTH completed AND non-completed homework. However, in discussing completed homework, the focus was MORE on the quantity of what was completed (i.e., the extent of completion), RATHER THAN the quality (i.e., degree of client learning or skill acquisition, such as mastery in completing a thought record effectively, or testing out beliefs in behavioral experiments).

6 BOTH the quantity (i.e., the extent of completion and non-completion) AND quality (i.e., degree of client learning or skill acquisition, such as mastery in completing a thought record effectively, or testing out beliefs in behavioral experiments) of homework completion was discussed. The therapist facilitated a highly effective discussion to elicit the CLIENT'S LEARNING from the homework task (e.g., using Socratic questioning).

0	1	2	3	4	5	6
Non-adherence/ extremely poor	Poor	Mediocre	Fair	Good	Very Good	Excellent

Item 2

2a **DID the therapist provide verbal reinforcement (i.e., praise) for any portion of the homework carried out?** Yes ☐ No ☐

2b **HOW WELL did the therapist provide appropriate verbal reinforcement (i.e., praise) for any portion of the homework carried out?** Competence Rating ☐

0 The therapist DID NOT provide verbal reinforcement for any portion of the homework carried out.

1 Verbal reinforcement was given that was VERY BRIEF AND LIMITED in relation to the portion of homework completed, OR excessive praise was given for low completion.

2 SOME verbal reinforcement was given but this was NOT CLEARLY LINKED to the portion of homework completed, OR excessive praise was given for low completion.

3 Appropriate verbal reinforcement was given for MOST portions of the homework completed.

4 Appropriate verbal reinforcement was given for ALL portions of the homework completed.

5 Appropriate praise AND encouragement was given for ALL portions of the homework completed. The therapist ALSO appeared clearly enthusiastic in acknowledging and validating the client's efforts.

6 Appropriate praise AND encouragement was given for ALL portions of the homework completed. The therapist ALSO appeared clearly enthusiastic in acknowledging and validating the client's efforts. Encouragement was given for the client EXTENDING/ GENERALIZING the homework task to extend skill acquisition/ apply task to more challenging problems.

0	1	2	3	4	5	6
Non-adherence/ extremely poor	Poor	Mediocre	Fair	Good	Very Good	Excellent

Item 3

3a **WAS a situational conceptualization (e.g., thoughts, behaviors, emotions, physiology) used in reviewing previously assigned homework?** Yes ☐ No ☐

3b **HOW WELL did the therapist use a situational conceptualization (e.g., thoughts, behaviors, emotions, physiology) to review previously assigned homework (i.e., identify the client's beliefs about having engaged in the homework to synthesize their learning)?** Competence Rating ☐

0 A situational conceptualization WAS NOT used in reviewing previously assigned homework.

1 An UNDEVELOPED situational conceptualization was arrived at (i.e., the therapist completely interpreted on behalf of the client).

2 A VAGUE, brief and incomplete situational conceptualization was arrived at (i.e., the therapist mostly interpreted for the client's experiences rather than eliciting information).

3 A PARTIALLY DEVELOPED situational conceptualization was arrived at (i.e., the therapist elicited some information and interpreted other information). NO automatic thoughts OR beliefs about the consequences, OR synthesis of learning were identified.

4 A situational conceptualization facilitated the IDENTIFICATION OF salient (i.e., emotionally laden) automatic thoughts, emotions, behaviors, and physiology that served as the TRIGGERS for homework completion.

5 A situational conceptualization facilitated the IDENTIFICATION OF salient (i.e., emotionally laden) automatic thoughts, emotions, behaviors, and physiology that served as the TRIGGERS for homework completion. The therapist ALSO elicited beliefs about the homework (i.e., difficulty, sense of pleasure, sense of mastery).

6 A situational conceptualization facilitated the IDENTIFICATION OF salient (i.e., emotionally laden) automatic thoughts, emotions, behaviors, and physiology that served as the TRIGGERS for homework completion. The therapist ALSO elicited beliefs about the homework (i.e., difficulty, sense of pleasure, sense of mastery) AS WELL AS their synthesis of learning (i.e., relevance, match with therapy goals, benefits, perceived progress)

0	1	2	3	4	5	6
Non-adherence/ extremely poor	Poor	Mediocre	Fair	Good	Very Good	Excellent

Item 4

4a **WAS an individualized conceptualization used to make sense of any portion of non-completed homework (i.e., linked non-completion to the client's automatic thoughts, underlying assumptions and rules, or core beliefs)?** Yes ☐ No ☐ n/a ☐

4b **HOW WELL did the therapist use an individualized conceptualization to make sense of any portion of non-completed homework (i.e., linked non-completion to the client's automatic thoughts, underlying assumptions and rules, or core beliefs)?** Competence Rating ☐

N/a There was NO NON-COMPLETED homework (i.e., all homework was completed).

0 The therapist DID NOT use an individualized conceptualization to make sense of any portion of non-completed homework.

1 The therapist LABELED/ INTERPRETED the portion of non-completed homework RATHER THAN facilitating the client's own understanding through collaborative discussion.

2 The therapist FOCUSED on one individualized conceptualization component (i.e., either core beliefs or conditional rules and assumptions or automatic thoughts). The therapist used this information to LABEL/ INTERPRET the portion of non-completed homework RATHER THAN facilitating the client's own understanding.

3 The therapist made LIMITED use of an individualized conceptualization, including SOME but NOT ALL of the following aspects: core beliefs, conditional rules and assumptions, and automatic thoughts. The therapist used this information to reach a VAGUE understanding of homework non-completion.

4 The therapist facilitated a discussion that made REASONABLE use of an individualized conceptualization, including SOME but NOT ALL aspects of: core beliefs, conditional rules and assumptions, and automatic thoughts. This led to a REASONABLE understanding of the client's beliefs about the homework task that contributed to non-completion.

5 The therapist facilitated a discussion that made GOOD USE of an individualized conceptualization including ALL ASPECTS of: core beliefs, conditional rules and assumptions, and automatic thoughts. This led to a CLEAR understanding of the client's beliefs about the homework task that contributed to non-completion.

6 The therapist facilitated a discussion that made FULL USE of an individualized conceptualization, including ALL ASPECTS of: core beliefs, conditional rules and assumptions, and automatic thoughts IN several situations, which were LINKED to overall treatment goals. This led to a VERY CLEAR understanding of the client's beliefs about the homework task that contributed to non-completion. AS WELL AS the generalization of the task to other situations.

0	1	2	3	4	5	6
Non-adherence/ extremely poor	Poor	Mediocre	Fair	Good	Very Good	Excellent

Item 5

5a **DID the therapist attempt to problem solve practical obstacles to the homework?** Yes ☐ No ☐

5b **HOW WELL did the therapist attempt to problem solve practical obstacles to the homework?** Competence Rating ☐

0 The therapist DID NOT attempt to problem solve practical obstacles.

1 The therapist PROVIDED solutions of their own accord, WITHOUT any contribution from the client.

2 The therapist PROVIDED solutions of their own accord, with only a CURSORY contribution sought from the client. (e.g., "Does that sound okay to you?").

3 The therapist ATTEMPTED to problem solve practical obstacles with SOME collaboration (i.e., the therapist provided some solutions themselves and elicited some input from the client).

4 The therapist FACILITATED a discussion that IDENTIFIED the actual practical obstacles. SOME potential solutions were generated and considered. The client arrived at a VAGUE plan to overcome the obstacles.

5 The therapist FACILITATED a discussion that IDENTIFIED the actual practical obstacles. A RANGE of potential solutions were generated and considered. The client arrived at CLEAR behavioral strategies to overcome the practical obstacles.

6 The therapist FACILITATED a discussion that IDENTIFIED the actual practical obstacles. AS WELL AS a consideration of other potential obstacles that may have occurred. A FULL RANGE of potential solutions were generated and considered. The client arrived at CLEAR behavioral strategies to overcome the practical obstacles. AS WELL AS behavioral strategies for considering changing circumstances (e.g., bringing an outside activity indoors, testing beliefs in several situations, applying interpersonal skills to a range of relationships/interactions).

OVERALL RATING: HOMEWORK REVIEW

0	1	2	3	4	5	6
Non-adherence/ extremely poor	Poor	Mediocre	Fair	Good	Very Good	Excellent

Please look over your ratings for items 1-5. Now provide one overall rating for HOMEWORK REVIEW . Please take into account:

- the individual ratings for items 1-5.
- the appropriateness of not adhering to specific items, e.g., homework was completed unusually well; there was a crisis or risk to client safety.
- any other special considerations from the session rated e.g., interpersonal features of the specific therapeutic relationship; and the therapist's ability to adapt the PROCESS AND DISCUSSION of homework based on the client's individualized cognitive conceptualization (e.g., greater verbal encouragement for a client with avoidant interpersonal style, normalizing aspects of non-completion for a client exhibiting perfectionism, emphasizing complimentary nature to existing coping strategies for client with demanding interpersonal style).

Rating

Please describe any factors that have affected your overall rating for HOMEWORK REVIEW:

HOMEWORK DESIGN Items 6-14 cover the therapist behaviors involved in collaboratively deciding what homework will be carried out between sessions, and typically occurs throughout the session.

0	1	2	3	4	5	6
Non-adherence/ extremely poor	Poor	Mediocre	Fair	Good	Very Good	Excellent

Item 6

6a **WAS any new or revised homework discussed?** Yes ☐ No ☐

6b **HOW WELL did the therapist discuss new or revised homework?** Competence Rating ☐

0 The therapist did NOT discuss new or revised homework.

1 The therapist BRIEFLY discussed new or revised homework.

2 The therapist allowed SUFFICIENT TIME for a discussion of new or revised homework, BUT only at the END of the session.

3 The therapist allowed SUFFICIENT TIME for a discussion of new or revised homework, throughout the course of the session. However, the homework WAS NOT linked to in-session content or therapy goals.

4 The therapist allowed SUFFICIENT TIME for a discussion of new or revised homework throughout the course of the session. The homework WAS linked to EITHER in-session content OR therapy goals.

5 The therapist allowed SUFFICIENT TIME for a discussion of new or revised homework throughout the course of the session. The homework WAS linked to BOTH in-session content AND therapy goals.

6 The therapist allowed SUFFICIENT TIME for a discussion of new or revised homework throughout the course of the session. The homework WAS linked to BOTH in-session content AND therapy goals. The therapist was ALSO able to tailor the discussion of the homework to the client's interpersonal style. The discussion was ALSO EFFECTIVE even when confronted with interpersonal difficulties (e.g., client avoidance, perfectionism, demanding interpersonal style).

0	1	2	3	4	5	6
Non-adherence/ extremely poor	Poor	Mediocre	Fair	Good	Very Good	Excellent

Item 7

Note:
This item asks about the therapist's use of the components of the "guided discovery" process. The guided discovery process has four sequential components which are:

i. Asking informational questions to uncover information outside the client's awareness,
ii. Listening empathically and providing reflections,
iii. Summarizing the information discovered,
iv. Asking synthesizing or analytical questions which enable the client's own learning

7a **DID the therapist use any aspects of guided discovery to identify the client's coping strategies and beliefs related to the homework?** Yes ☐ No ☐

7b **HOW WELL did the therapist use guided discovery to identify the client's coping strategies and beliefs related to the homework?** Competence Rating ☐

0 The therapist DID NOT use any aspects of guided discovery to identify the client's coping strategies and beliefs related to the homework.

1 The therapist used INEFFECTIVE questioning (e.g., closed questions or broad questions but these did not uncover new information) and provided INTERPRETIVE answers RATHER THAN guiding the client's own understanding about coping strategies and beliefs.

2 The therapist used SOME but NOT ALL components of the guided discovery process. HOWEVER they were used in a cursory, inappropriate, or ineffective manner (e.g., inaccurate reflections or summaries). The therapist used INTERPRETIVE answers RATHER THAN guiding the client's own learning, and was UNABLE to identify coping strategies and beliefs (e.g., "If you think X, then surely Y is ....?").

3 The therapist used ALL FOUR components of the guided discovery process BUT was INEFFECTIVE in identifying coping strategies and beliefs.

4 The therapist used ALL FOUR components of the guided discovery process REASONABLY EFFECTIVELY. In using this process the therapist facilitated the identification of A FEW coping strategies and beliefs.

5 The therapist used ALL FOUR components of the guided discovery process EFFECTIVELY. In using this process the therapist facilitated the identification of A NUMBER OF coping strategies and beliefs.

6 The therapist APPEARED genuinely curious and inquisitive, and used ALL FOUR components of the guided discovery process VERY EFFECTIVELY. In using this process the therapist facilitated the identification of A NUMBER OF HIGHLY CREDIBLE coping strategies and beliefs.

0	1	2	3	4	5	6
Non-adherence/ extremely poor	Poor	Mediocre	Fair	Good	Very Good	Excellent

Item 8

8a **DID the therapist integrate a disorder-specific cognitive model with the individualized conceptualization in designing homework?** Yes ☐ No ☐

8b **HOW WELL did the therapist integrate a disorder-specific cognitive model with the individualized conceptualization in designing homework?** Competence Rating ☐

0 The therapist DID NOT discuss a disorder-specific cognitive model or individualized conceptualization in designing homework.

1 The therapist MENTIONED the disorder-specific cognitive model BUT did not elaborate on how it was relevant to the client's presentation.

2 The therapist integrated SOME but NOT ALL aspects of a disorder-specific cognitive model (e.g., Beck's Cognitive Triad, Clark's Panic Model, etc) to ONE ASPECT of the client's individualized conceptualization (i.e., core beliefs, conditional rules and assumptions, automatic thoughts, and under and over developed behavioral strategies).

3 The therapist integrated SOME but NOT ALL aspects of a disorder-specific cognitive model (e.g., Beck's Cognitive Triad, Clark's Panic Model, etc) to MORE THAN ONE aspect the client's individualized conceptualization (i.e., core beliefs, conditional rules and assumptions, automatic thoughts, and under and over developed behavioral strategies).

4 The therapist integrated MOST aspects of a disorder-specific cognitive model (e.g., Beck's Cognitive Triad, Clark's Panic Model, etc) to MOST aspects the client's individualized conceptualization (i.e., core beliefs, conditional rules and assumptions, automatic thoughts, and under and over developed behavioral strategies).

5 The therapist integrated ALL aspects of a disorder-specific cognitive model (e.g., Beck's Cognitive Triad, Clark's Panic Model, etc) to ALL aspects of the client's individualized conceptualization (i.e., core beliefs, conditional rules and assumptions, automatic thoughts, and under and over developed behavioral strategies).

6 The therapist integrated ALL aspects of a disorder-specific cognitive model (e.g., Beck's Cognitive Triad, Clark's Panic Model, etc) to ALL aspects of the client's individualized conceptualization (e.g., core beliefs, conditional rules and assumptions, automatic thoughts, and under and over developed behavioral strategies). The therapist was ALSO able to integrate all this information with the client's presenting problems, evidenced in tactful responses to client's interpersonal style (e.g., critical, competitive, suspicious, controlling exaggerative).

0	1	2	3	4	5	6
Non-adherence/ extremely poor	Poor	Mediocre	Fair	Good	Very Good	Excellent

Item 9

9a **WERE homework tasks selected for completion before the next session?** Yes ☐ No ☐

9b **HOW WELL did the therapist collaboratively select homework tasks for completion before the next session?** Competence Rating ☐

0 Homework tasks were NOT selected during the session.

1 The therapist selected homework tasks WITHOUT any contribution of the client.

2 The therapist only sought a CURSORY contribution from the client in selecting homework tasks (e.g., "Does that sound okay to you?").

3 The therapist INVOLVED the client in the selection of homework tasks, BUT at times reverted to a DIRECTIVE rather than collaborative approach, especially in the final decision.

4 The therapist INVOLVED the client in the selection of homework tasks (E.G., facilitated a DISCUSSION rather than provided direct answers). A FEW possible homework tasks were discussed, AS WELL AS a FEW advantages and disadvantages of the possible homework tasks.

5 The therapist encouraged the client to view the process of selecting homework tasks as the therapist and client working together as a TEAM. The therapist also ACTIVELY INVOLVED the client in selecting homework tasks (e.g., facilitated a discussion rather than provided direct answers). SEVERAL possible homework tasks were discussed, AND the client's thoughts and feelings about the possible homework tasks were elicited and explored, AND SEVERAL advantages and disadvantages of the possible homework tasks were discussed.

6 The therapist encouraged the client to view the process of selecting homework tasks as the therapist and client working together as a TEAM. The therapist also ACTIVELY INVOLVED the client in selecting homework tasks (e.g., facilitated a discussion rather than provided direct answers). A FULL RANGE of possible homework tasks were discussed, AND the client's thoughts and feelings about the possible homework tasks were elicited and explored, AND A FULL RANGE of advantages and disadvantages of the possible homework tasks were discussed (i.e., based on prior experience, benefits experienced by others). The therapist and client ALSO decided on homework tasks that built upon existing client skills and strategies, AND the client was encouraged to take on more responsibility for selecting homework tasks.

0	1	2	3	4	5	6
Non-adherence/ extremely poor	Poor	Mediocre	Fair	Good	Very Good	Excellent

Item 10

10a **DID the therapist present any rationale for the homework?** Yes ☐ No ☐

10b **HOW WELL did the therapist present a rationale for the homework that aligned with the client's goals for treatment?** Rating ☐

0 The therapist DID NOT present any rationale for the homework.

1 The therapist presented a BRIEF rationale but FAILED to relate it to the client's treatment goals.

2 The therapist presented a RATIONALE for the homework with SOME mention of the client's treatment goals, however this was presented WITHOUT any input (and understanding) from the client.

3 The therapist assisted the client to understand how the homework was ALIGNED to the specific presenting problem in the current session.

4 The therapist assisted the client to understand that the homework was broken into achievable CHUNKS that were manageable and within the client's control. The therapist ALSO assisted the client to understand how the homework was ALIGNED to EITHER the specific presenting problem in the current session, OR their overall treatment goals.

5 The therapist assisted the client to understand that the homework was broken into achievable CHUNKS that were manageable and within the client's control. The therapist ALSO assisted the client to understand how the homework was ALIGNED to BOTH the specific presenting problem in the current session AS WELL AS their overall treatment goals.

6 The therapist assisted the client to understand that the homework was broken into achievable CHUNKS that were manageable and within the client's control. The therapist ALSO assisted the client to understand how the homework was ALIGNED to BOTH the specific presenting problem in the current session AS WELL AS their overall treatment goals, AND obtained feedback from the client on the rationale. The therapist ALSO provided empirical evidence to support the rationale for the homework.

0	1	2	3	4	5	6
Non-adherence/ extremely poor	Poor	Mediocre	Fair	Good	Very Good	Excellent

Item 11

11a DID the therapist ask about the client's ability and perceived difficulty of the homework? Yes ☐ No ☐

11b HOW WELL did the therapist ask about the client's ability and perceived difficulty of the homework? Rating ☐

0 The therapist DID NOT ask about the client's ability and perceived difficulty of the task

1 The therapist made a CURSORY enquiry about the client's ability and perceived difficulty of the task, but did not discuss it any further.

2 The therapist ENQUIRED about the client's ability and perceived difficulty of the task, and made an INEFFECTIVE attempt to elicit feedback from the client (e.g., the therapist did not listen to the client's responses, asked closed questions, questions did not follow the client's responses).

3 The therapist ENQUIRED about the client's ability and perceived difficulty of the task, and elicited a GENERAL STATEMENT from the client, for example, the client was vague and said "Sure, I can do it" and this response was taken at face value and NOT explored any further.

4 The therapist ENQUIRED about the client's ability and perceived difficulty of the task, and through Socratic questioning identified a BROAD ISSUE (e.g., "That thought record looks too hard. There is so much to complete"). HOWEVER, the therapist then provided their own solutions to resolve the issues raised (e.g., "Okay, just complete the first three columns of the thought record").

5 The therapist ENQUIRED about the client's ability and perceived difficulty of the task, and through Socratic questioning identified SPECIFIC ISSUES (e.g., in addition to feeling overwhelmed by the entire thought record, it transpired that the client had difficulty distinguishing emotions and thoughts on thought record). Through further EXPLORATION the therapist and client collaboratively RESOLVED the issue (e.g., the therapist and client worked on automatic thoughts in-session, and/or the homework was redesigned to focus on practicing the identification emotions as distinct from automatic thoughts).

6 The therapist ENQUIRED about the client's ability and perceived difficulty of the task, and through Socratic questioning identified SPECIFIC ISSUES. Through further EXPLORATION the therapist and client collaboratively RESOLVED the issue. The therapist ALSO elicited ADDITIONAL CLIENT LEARNING from the discussion, for example, the client learnt that breaking items into smaller chunks was less overwhelming, and also identified an underlying rule (e.g., "I've failed if I can't work things out for myself").

0	1	2	3	4	5	6
Non-adherence/ extremely poor	Poor	Mediocre	Fair	Good	Very Good	Excellent

Item 12

12a **WAS ANY attempt made to facilitate in-session homework practice?** Yes ☐ No ☐

12b **HOW WELL did the therapist facilitate in-session homework practice?** Rating ☐

0 The therapist DID NOT provide the opportunity for in-session practice of the homework.

1 The therapist briefly DEMONSTRATED or EXPLAINED (i.e., modeled or instructed) the homework, that provided no opportunity for the clients to learn from their own practice.

2 The therapist PROVIDED only a BRIEF opportunity for in-session practice. The therapist tended to FOCUS on correcting the client's mistakes AND provided LIMITED positive reinforcement. The therapist DID NOT discuss any learning points from the practice.

3 The therapist PROVIDED SOME opportunity for in-session practice. The therapist provided SOME positive reinforcement (i.e., shaping successive approximations of skill). AND gave SOME constructive guidance when the client needed assistance. HOWEVER, the therapist used a DIRECTIVE rather than collaborative approach in discussing learning points from the practice.

4 The therapist PROVIDED SOME opportunity for in-session practice. The therapist provided SOME positive reinforcement (i.e., shaping successive approximations of skill). AND gave SOME constructive guidance when the client needed assistance. The therapist and client COLLABORATIVELY discussed learning points from the in-session practice.

5 The therapist PROVIDED a GOOD opportunity for in-session practice, using the METHOD/S most appropriate for the client and the specific task. The therapist provided POSITIVE reinforcement (i.e., shaping successive approximations of skill) AND gave CONSTRUCTIVE guidance when the client needed assistance. The therapist was ENCOURAGING when COLLABORATIVELY discussing learning points from the in-session practice.

6 The therapist PROVIDED CONSIDERABLE opportunity for in-session practice, using the METHOD/S most appropriate for the client and the specific task. The therapist provided ENTHUSIASTIC positive reinforcement (i.e., shaping successive approximations of skill), AND gave WARM, GENUINE, CONSTRUCTIVE guidance when the client needed assistance. The therapist was ENCOURAGING when COLLABORATIVELY discussing learning points from the in-session practice. The therapist ALSO asked the client for FEEDBACK on the experience, and asked the client to WRITE down the learning points.

0	1	2	3	4	5	6
Non-adherence/ extremely poor	Poor	Mediocre	Fair	Good	Very Good	Excellent

Item 13

13a DID the therapist use guided imagery to begin experiential learning for the homework in-session? Yes ☐ No ☐

13b HOW WELL did the therapist use guided imagery to begin experiential learning for the homework in-session? Rating ☐

0 The therapist DID NOT use guided imagery in homework design.

1 The therapist used guided imagery INEFFECTIVELY (i.e., affect was not generated, client had difficulty staying on track, etc). Feedback was NOT sought throughout the exercise, and at the completion, the therapist DID NOT facilitate any experiential learning from the imagery practice.

2 The therapist provided an OPPORTUNITY for guided imagery , but was UNABLE to use this to assist the client with some experiential learning of the homework task (i.e., client completed imagery but did not gain an experience of completing the task).

3 The therapist FACILITATED the client in using guided imagery, and this was REASONABLY EFFECTIVE in stepping the client through a scenario where they may use the homework assignment (i.e., physiological, emotional, cognitive triggers identified). HOWEVER, imagery was INEFFECTIVE in providing the client with some experiential learning of the homework task (i.e., client completed imagery but did not gain an experience of completing the task).

4 The therapist FACILITATED the client in using guided imagery, and this was REASONABLY EFFECTIVE in stepping the client through a scenario where they may use the homework assignment (i.e., physiological, emotional, cognitive triggers identified), AND the client gained SOME experiential learning of the homework task (i.e., experienced the outcome of having engaged in the homework task).

5 The therapist FACILITATED the client in using guided imagery, and this was EFFECTIVE in stepping the client through a scenario where they may use the homework assignment (i.e., physiological, emotional, cognitive triggers identified), AND the client gained SOME experiential learning of the homework task (i.e., experienced the outcome of having engaged in the homework task). The therapist ALSO focused on skill acquisition AND discussed with the client how the task could be extended to more complex skills (i.e., shaping).

6 The therapist FACILITATED the client in using guided imagery, and this was EFFECTIVE in stepping the client through a scenario where they may use the homework assignment (i.e., physiological, emotional, cognitive triggers identified), AND the client gained SOME experiential learning of the homework task (i.e., experienced the outcome of having engaged in the homework task). The therapist ALSO focused on skill acquisition AND discussed with the client how the task could be extended to more complex skills (i.e., shaping). In feedback, the therapist and client ALSO discussed the application of the task across different situations (i.e., generalization and maintenance).

0	1	2	3	4	5	6
Non-adherence/ extremely poor	Poor	Mediocre	Fair	Good	Very Good	Excellent

Item 14

14a **DID the therapist use a situational conceptualization to help identify the client's beliefs and triggers (i.e., emotional, behavioral, physiological) for carrying out the homework in specific situations?** Yes ☐ No ☐

14b **HOW WELL did the therapist use a situational conceptualization to help identify the client's beliefs and triggers (i.e., emotional, behavioral, physiological) for carrying out the homework in specific situations?** Rating ☐

0 The therapist DID NOT use a situational conceptualization to help identify the client's beliefs and situational triggers for carrying out the homework in specific situations.

1 An UNDEVELOPED situational conceptualization was arrived at (i.e., the therapist completely interpreted on behalf of the client).

2 A VAGUE, BRIEF AND INCOMPLETE situational conceptualization was arrived at (i.e., the therapist mostly interpreted for the client rather than eliciting information).

3 A PARTIALLY DEVELOPED situational conceptualization was arrived at (i.e., the therapist elicited some information and interpreted other information). This PROVED INEFFECTIVE in identifying the client's beliefs and situational triggers.

4 A PARTIALLY DEVELOPED situational conceptualization was arrived at (i.e., the therapist elicited some information and interpreted other information). Emotions, behaviors, and physiology WERE IDENTIFIED to the use of homework. BUT no cognitive triggers or beliefs were identified.

5 A SITUATIONAL CONCEPTUALIZATION facilitated the client's identification of SALIENT (i.e., emotionally laden) automatic thoughts that served as triggers for homework completion. Emotions, behaviors, and physiology were also identified.

6 A SITUATIONAL CONCEPTUALIZATION facilitated the client's identification of a SALIENT (i.e., emotionally laden) automatic thoughts, emotions, behaviors, and physiology that served as triggers for homework completion. The therapist ALSO discussed the triggers to the use of homework in several situations. AND elicited beliefs about the homework (i.e., difficulty, obstacles).

OVERALL RATING: HOMEWORK DESIGN

0	1	2	3	4	5	6
Non-adherence/ extremely poor	Poor	Mediocre	Fair	Good	Very Good	Excellent

Please look over your ratings for items 6-14. Now provide one overall rating for HOMEWORK DESIGN . Please take into account:

- the individual ratings for items 6-14.
- the appropriateness of not adhering to specific items, e.g., no need to practice a particular skill for homework as this had covered extensively in previous sessions; client was extending a mastered skill to a new situation rather than being asked to learn something new.
- and any other special considerations from the session rated, e.g., interpersonal features of the specific therapeutic relationship, and the therapist's ability to adapt the PROCESS AND DISCUSSION of homework based on the client's individualized cognitive conceptualization (e.g., increased emphasis on in-session practice for a client with dependent interpersonal style, discussion of rationale for a client with controlling interpersonal style).

Rating

Please describe any factors that have affected your overall rating for HOMEWORK DESIGN.

HOMEWORK ASSIGN Items 15-19 cover the therapist behaviors involved in determining how the homework assignment/s will be practically carried out, and typically occurs near the end of the session.

0	1	2	3	4	5	6
Non-adherence/ extremely poor	Poor	Mediocre	Fair	Good	Very Good	Excellent

Item 15

15a **WAS there any attempt to summarize the rationale for the homework in relation to therapy goals?** Yes ☐ No ☐

15b **HOW WELL did the therapist ask the client to summarize the rationale for the homework in relation to therapy goals?** Rating ☐

0 The therapist DID NOT ask the client to summarize the rationale for the task in relation to therapy goals.

1 The therapist summarized the rationale for the task, with LITTLE OR NO INPUT from the client.

2 The therapist ATTEMPTED to involve the client in summarizing the rationale for the task in relation to therapy goals, but used a DIRECTIVE rather than collaborative approach.

3 The therapist INVOLVED the client in summarizing the rationale for the task in relation to GENERAL therapy goals.

4 The therapist INVOLVED the client in summarizing the rationale for the task in relation to MOST PERTINENT therapy goals. That is, the homework was discussed in terms of the SPECIFIC behavior changes that would be expected to result from progress towards this goal.

5 The therapist SKILLFULLY INVOLVED the client in summarizing the rationale for the task in relation to MOST PERTINENT therapy goals. That is, the homework was discussed in terms of the SPECIFIC behavior changes that would be expected to result from progress towards this goal, AND this process was LEAD by the client.

6 The therapist SKILLFULLY INVOLVED the client in summarizing the rationale for the task in relation to MOST PERTINENT therapy goals. That is, the homework was discussed in terms of the SPECIFIC behavior changes that would be expected to result from progress towards this goal, AND this process was LEAD by the client. FURTHERMORE, in discussion with the therapist, the client demonstrated a clear understanding of the homework and was able to place the current homework in context of current and overall goals for therapy. The therapist skill was evidenced by their adaptation of this discussion to the client's interpersonal style.

0	1	2	3	4	5	6
Non-adherence/ extremely poor	Poor	Mediocre	Fair	Good	Very Good	Excellent

Item 16

16a **WAS there any attempt to specify how the homework will be practically integrated into the client's life (i.e., specification of when, where, how often, how long)?** Yes ☐ No ☐

16b **HOW WELL did the therapist collaborate with the client to specify how the homework will be practically integrated into the client's life (i.e., specification of when, where, how often, how long)?** Rating ☐

0 The therapist DID NOT collaborate to specify how the task would be practically integrated into the client's life.

1 The therapist DIRECTED how the task could be practically integrated into the client's life, WITHOUT any contribution from the client.

2 The therapist reached a VAGUE outline of how the task could be practically integrated into the client's life, with SOME collaboration (i.e., the therapist provided some specifics themselves and elicited some input from the client).

3 The therapist FACILITATED a discussion which resulted in the client being able to state with SOME behavioral specificity how the task could be practically integrated into the client's life in ONE of the following areas: when, where, how often, and how long

4 The therapist FACILITATED a discussion which resulted in the client being able to state with SOME behavioral specificity how the task could be practically integrated into the client's life in TWO-THREE of the following areas: when, where, how often, and how long.

5 The therapist FACILITATED a discussion which resulted in the client being able to state with a HIGH DEGREE of behavioral specificity how the task could be practically integrated into the client's life in ALL of the following areas: when, where, how often, and how long. IF the client was unable to be specific in any area, the therapist gently GUIDED the client to a specific resolution.

6 The therapist SKILLFULLY ELICITED a description of how the homework would be practically implemented from the client. A HIGH DEGREE of behavioral specificity was achieved in ALL the following areas: when, where, how often, and how long. IF the client was unable to be specific in any area, the therapist gently GUIDED the client to a specific resolution. The therapist ALSO anticipated potential difficulties in communication and resolved them (e.g., misinterpretation of the process in achieving specificity, misinterpretation of the meaning of specificity, such as using a thought record "when" automatic thoughts occur).

0	1	2	3	4	5	6
Non-adherence/ extremely poor	Poor	Mediocre	Fair	Good	Very Good	Excellent

Item 17

17a **WAS there any consideration of potential difficulties for completing the homework?** Yes ☐ No ☐

17b **HOW WELL did the therapist consider potential difficulties for completing the homework?** Rating ☐

0 The therapist DID NOT attempt to consider potential difficulties.

1 The therapist PROVIDED potential difficulties of their own accord, WITHOUT any contribution from the client.

2 The therapist GENERALLY PROVIDED potential difficulties of their own accord, with only a CURSORY CONTRIBUTION sought from the client. (e.g., "So that would be difficult, wouldn't it?").

3 The therapist attempted to consider potential difficulties with some collaboration (i.e., the therapist provided some potential difficulties themselves and elicited some input from the client).

4 The therapist FACILITATED a discussion that identified SOME potential difficulties, AND SOME potential solutions were also generated and considered. The client arrived at a VAGUE plan to overcome the potential difficulties.

5 The therapist FACILITATED a discussion that identified MOST potential difficulties, AND a RANGE of potential solutions were generated and considered. The client arrived at a CLEAR plan to overcome the potential difficulties that included SPECIFIC behaviors (e.g., "My days are really busy next week, so I will set the alarm clock 30 minutes earlier on Tuesday morning and read the booklet before starting the day's other activities").

6 The therapist FACILITATED a discussion that identified ALL the potential difficulties, and a FULL RANGE of potential solutions were generated and considered. The client arrived at a CLEAR plan to overcome the potential difficulties that included SPECIFIC behaviors, AND behavioral STRATEGIES for considering changing circumstances (e.g., if unable to complete a task in a single sitting, then breaking it into smaller chunks and completing it over 2-3 sittings).

0	1	2	3	4	5	6
Non-adherence/ extremely poor	Poor	Mediocre	Fair	Good	Very Good	Excellent

Item 18

18a **WAS there ANY attempt to explain the outcome from the homework as having a learning 'experiment' focus?** Yes ☐ No ☐

18b **HOW WELL did the therapist emphasize the homework as having a learning 'experiment' focus (e.g., a no-lose scenario, partial completion is helpful, seeing what works and what doesn't)?** Rating ☐

0 The therapist DID NOT emphasize the task as a learning 'experiment' focus.

1 In ATTEMPTING to explain a learning 'experiment' focus of the homework task, the therapist specified or intimated there was a 'CORRECT' actual outcome (i.e. could pass or fail).

2 The therapist did not focus on actual outcomes, but was VAGUE about the learning outcome (i.e., "it will be useful") but did not elaborate any further.

3 The therapist BRIEFLY explained the homework task as a learning experiment (i.e., to test out an idea or skill), rather than guided the client to their own learning.

4 The therapist FRAMED the homework task as a learning 'experiment'. MOST of the following points emerged from the discussion: there is no right or wrong (no failure or grading); it is a no-lose situation for the client; in any experiment the outcome is not known; there is a learning from every homework task no matter what the actual outcome; any information from the experiment is useful to further help with the treatment.

5 The therapist used guided discovery to uncover the CLIENT'S BELIEFS about the outcomes of the homework task, and then used Socratic questioning and hypothetical examples to facilitate the CLIENT to view the homework task as a learning experiment (i.e., gaining client's previous experiences of learning and applying them to the homework). MOST of the following points emerged from the discussion: there is no right or wrong (no failure or grading); it is a no-lose situation for the client; in any experiment the outcome is not known; there is a learning from every homework task no matter what the actual outcome; any information from the experiment is useful to further help with the treatment.

6 The therapist used guided discovery to uncover the CLIENT'S BELIEFS about the outcomes of the homework task, and then used Socratic questioning and hypothetical examples to facilitate the CLIENT to view the homework task as a learning experiment(i.e., gaining client's previous experiences of learning and applying them to the homework). MOST of the following points emerged from the discussion: there is no right or wrong (no failure or grading); it is a no-lose situation for the client; in any experiment the outcome is not known; there is a learning from every homework task no matter what the actual outcome; any information from the experiment is useful to further help with the treatment. The therapist ALSO discussed the BENEFITS (e.g., new skill acquisition, reduction in distressing thoughts, better treatment outcome) VERSUS the COSTS of performing the homework task (e.g., time, energy, short-term distress).

0	1	2	3	4	5	6
Non-adherence/ extremely poor	Poor	Mediocre	Fair	Good	Very Good	Excellent

Item 19

19a **WAS there ANY attempt to summarize the homework?** Yes ☐ No ☐

19b **HOW WELL did the therapist ask the client to summarize the homework and obtain an indication of homework-related readiness, importance, and/ or confidence?** Rating ☐

0 There was NO summary of the homework task AND NO indication of readiness, importance, or confidence.

1 The therapist summarized the task, WITHOUT any contribution from the client , AND DID NOT obtain any indication of readiness, importance, or confidence.

2 The therapist ATTEMPTED to involve the client in summarizing the task AND obtained separate indications for readiness, importance, or confidence, with only a CURSORY contribution sought from the client. (e.g., "Does about 80% sound right to you?").

3 The therapist INVOLVED the client in summarizing the task and obtained a VAGUE indication of readiness, importance, or confidence (e.g., the client said "I'd give that a very high rating").

4 The therapist FACILITATED the client to SUMMARIZE the task AND provide an indication of readiness, importance, and confidence.

5 The therapist USED Socratic questioning, which enabled the client to SUMMARIZE the task AND provide SPECIFIC ratings for EACH OF readiness, importance, or confidence. IF the task summary was incomplete, the client was gently guided to its completion. IF the rating was low (i.e., <70%) the client was gently guided to identify what it would take to increase their rating.

6 The therapist USED Socratic questioning, which enabled the client to ACTIVELY SUMMARIZE the task AND provide SPECIFIC ratings for EACH OF readiness, importance, and confidence. IF the task summary was incomplete, the client was gently guided to its completion. OR the task was modified with decreased demands. IF the confidence rating was low (i.e., <70%) the client was gently guided to identify what it would take to increase their confidence level. The therapist ALSO explored overly confident ratings (e.g., an immediate or persistent statement of 100%) to identify possible social desirability responses.

OVERALL RATING: HOMEWORK ASSIGN

0	1	2	3	4	5	6
Non-adherence/ extremely poor	Poor	Mediocre	Fair	Good	Very Good	Excellent

Please look over your ratings for items 15-19. Now provide one overall rating for HOMEWORK ASSIGN. Please take into account:

- the individual ratings for items 15-19
- the appropriateness of not adhering to specific items (e.g., the session was near the end of therapy and the client was taking responsibility for the process, and leading in-session discussion)
- and any other special considerations from the session rated, e.g., interpersonal features of the specific therapeutic relationship; and the therapist's ability to adapt the PROCESS AND DISCUSSION of homework based on the client's individualized cognitive conceptualization (e.g., greater use of collaboration for a client with suspicious interpersonal style, greater discussion of specifics and confidence ratings for a client with unrelenting standards).

Rating

Please describe any factors that have affected your overall rating for HOMEWORK ASSIGN:

APPENDIX B

THERAPIST HOMEWORK ADHERENCE AND COMPETENCE SCALE

Therapist Homework Assignment Competency Scale* **Directions:** For each of the following items, note the description of relevant therapist behavior and put a check (√) in the line column to rate the therapist on the scale, from 0 to 4. **Please rate all the items.**	**0--Very Poor or not Done**	**1--Somewhat less than Satisfactory**	**2--Satisfactory**	**3--Somewhat more than Satisfactory**	**4--Very Well Done**
1. REVIEWING PREVIOUS SESSIONS HOMEWORK Near the beginning of each session, barring problems needing immediate attention, the therapist should review homework assigned the preceding session, and summarize progress made and conclusions drawn from the exercise.					
2. PROVIDING A RATIONALE The therapist stresses not only the importance of the assignment, but also the goals it aims to accomplish.					
3. ASSIGNING HOMEWORK The therapist should be clear and specific in giving assignments. The assignments should be "custom-tailored" to the individual client based on problems discussed during the session.					
4. ELICITING REACTIONS AND POSSIBLE DIFFICULTIES The therapist should elicit reactions to assignments to get feedback and help foresee and problem-solve possible obstacles or difficulties that might arise from trying to do the assignment. If this is a session in the middle or later in the treatment, the therapist should encourage the client to suggest or help design the assignment.					

*** (THACS) Bryant, Simons, and Thase (1998)**

APPENDIX C

COGNITIVE THERAPY SCALE

Cognitive Therapy Scale

Therapist:_____ Patient:_____ Date of Session:_____

Tape ID#:_____ Rater:_____ Date of Rating:_____

Session#_____ ( ) Videotape ( ) Audiotape ( ) Live Observation

Directions: For each time, assess the therapist on a scale from 0 to 6, and record the rating on the line next to the item number. Descriptions are provided for even-numbered scale points. If you believe the therapist falls between two of the descriptors, select the intervening odd number (1, 3, 5). For example, if the therapist set a very good agenda but did not establish priorities, assign a rating of a 5 rather than a 4 or 6.

If the descriptions for a given item occasionally do not seem to apply to the session you are rating, feel free to disregard them and use the more general scale below:

0	1	2	3	4	5	6
Poor	Barely Adequate	Mediocre	Satisfactory	Good	Very Good	Excellent

Please do not leave any item blank. For all items, focus on the skill of the therapist, taking into account how difficult the patient seems to be.

Part I. GENERAL THERAPEUTIC SKILLS

___1. AGENDA

0 Therapist did not set agenda.

2 Therapist set agenda that was vague or incomplete.

4 Therapist worked with patient to set a mutually satisfactory agenda that included specific target problems (e.g., anxiety at work, dissatisfaction with marriage.)

6 Therapist worked with patient to set an appropriate agenda with target problems, suitable for the available time. Established priorities and then followed agenda.

___2. FEEDBACK

0 Therapist did not ask for feedback to determine patient's understanding of, or response to, the session.

2 Therapist elicited some feedback from the patient, but did not ask enough questions to be sure the patient understood the therapist's line of reasoning during the session or to ascertain whether the patient was satisfied with the session.

4 Therapist asked enough questions to be sure that the patient understood the therapist's line of reasoning throughout the session and to determine the patient's reactions to the session. The therapist adjusted his/her behavior in response to the feedback, when appropriate.

6 Therapist was especially adept at eliciting and responding to verbal and non-verbal feedback throughout the session (e.g., elicited reactions to session, regularly checked for understanding, helped summarize main points at end of session.

___3. UNDERSTANDING

0 Therapist repeatedly failed to understand what the patient explicitly said and thus consistently missed the point. Poor empathic skills.

2 Therapist was usually able to reflect or rephrase what the patient explicitly said, but repeatedly failed to respond to more subtle communication. Limited ability to listen and empathize.

4 Therapist generally seemed to grasp the patient's "internal reality" as reflected by both what the explicitly said and what the patient communicated in more subtle ways. Good ability to listen and empathize.

6 Therapist seemed to understand the patient's "internal reality" thoroughly and was adept at communicating this understanding through appropriate verbal and non-verbal responses to the patient (e.g., the tone of the therapist's response conveyed a sympathetic understanding of the patient's "message"). Excellent listening and empathic skills.

___4. INTERPERSONAL EFFECTIVENESS

0 Therapist had poor interpersonal skills. Seemed hostile, demeaning, or in some other way destructive to the patient.

2 Therapist did not seem destructive, but had significant interpersonal problems. At times, therapist appeared unnecessarily impatient, aloof, insincere or had difficulty conveying confidence and competence.

4 Therapist displayed a satisfactory degree of warmth, concern, confidence, genuineness, and professionalism. No significant interpersonal problems.

6 Therapist displayed optimal levels of warmth, concern, confidence, genuineness, and professionalism, appropriate for this particular patient in this session.

___5. COLLABORATION

0 Therapist did not attempt to set up a collaboration with patient.

2 Therapist attempted to collaborate with patient, but had difficulty either defining a problem that the patient considered important or establishing rapport.

4 Therapist was able to collaborate with patient, focus on a problem that both patient and therapist considered important, and establish rapport.

6 Collaboration seemed excellent; therapist encouraged patient as much as possible to take an active role during the session (e.g., by offering choices) so they could function as a "team".

___6. PACING AND EFFICIENT USE OF TIME

0 Therapist made no attempt to structure therapy time. Session seemed aimless.

2 Session had some direction, but the therapist had significant problems with structuring or pacir (e.g., too little structure, inflexible about structure, too slowly paced, too rapidly paced).

4 Therapist was reasonably successful at using time efficiently. Therapist maintained appropria control over flow of discussion and pacing.

6 Therapist used time efficiently by tactfully limiting peripheral and unproductive discussion and t pacing the session as rapidly as was appropriate for the patient.

Part II. CONCEPTUALIZATION, STRATEGY, AND TECHNIQUE

___7. GUIDED DISCOVERY

0 Therapist relied primarily on debate, persuasion, or "lecturing". Therapist seemed to be "cros examining" patient, putting the patient on the defensive, or forcing his/her point of view on tl patient.

2 Therapist relied too heavily on persuasion and debate, rather than guided discovery. Howeve therapist's style was supportive enough that patient did not seem to feel attacked or defensive.

4 Therapist, for the most part, helped patient see new perspectives through guided discovery (e.g examining evidence, considering alternatives, weighing advantages and disadvantages) rather tha through debate. Used questioning appropriately.

6 Therapist was especially adept at using guided discovery during the session to explore problen and help patient draw his/her own conclusions. Achieved an excellent balance between skillf questioning and other modes of intervention.

___8. FOCUSING ON KEY COGNITIONS OR BEHAVIORS

0 Therapist did not attempt to elicit specific thoughts, assumptions, images, meanings, behaviors.

2 Therapist used appropriate techniques to elicit cognitions or behaviors; however, therapist ha difficulty finding a focus or focused on cognitions/behaviors that were irrelevant to the patient's ke problems.

4 Therapist focused on specific cognitions or behaviors relevant to the target problem. Howeve therapist could have focused on more central cognitions or behaviors that offered greater promise fo progress.

6 Therapist very skillfully focused on key thoughts, assumptions, behaviors, etc. th were most relevant to the problem area and offered considerable promise for progress.

___9. STRATEGY FOR CHANGE (Note: For this item, focus on the quality of the therapist's strategy for change, not on how effectively the strategy was implemented or whether change actually occurred.)

0 Therapist did not select cognitive-behavioral techniques.

2 Therapist selected cognitive-behavioral techniques; however, either the overall strategy for bringing about change seemed vague or did not seem promising in helping the patient.

4 Therapist seemed to have a generally coherent strategy for change that showed reasonable promise and incorporated cognitive-behavioral techniques.

6 Therapist followed a consistent strategy for change that seemed very promising and incorporated the most appropriate cognitive-behavioral techniques.

___10. APPLICATION OF COGNITIVE-BEHAVIORAL TECHNIQUES (Note: For this item, focus on how skillfully the techniques were applied, not on how appropriate they were for the target problem or whether change actually occurred.)

0 Therapist did not apply any cognitive-behavioral techniques.

2 Therapist used cognitive-behavioral techniques, but there were significant flaws in the way they were applied.

4 Therapist applied cognitive-behavioral techniques with moderate skill.

6 Therapist very skillfully and resourcefully employed cognitive-behavioral techniques.

___11. HOMEWORK

0 Therapist did not attempt to incorporate homework relevant to cognitive therapy.

2 Therapist had significant difficulties incorporating homework (e.g., did not review previous homework, did not explain homework in sufficient detail, assigned inappropriate homework).

4 Therapist reviewed previous homework and assigned "standard" cognitive therapy homework generally relevant to issues dealt with in session. Homework was explained in sufficient detail.

6 Therapist reviewed previous homework and carefully assigned homework drawn from cognitive therapy for the coming week. Assignment seemed "custom tailored" to help patient incorporate new perspectives, test hypotheses, experiment with new behaviors discussed during session, etc.

Part III. ADDITIONAL CONSIDERATIONS

12. (a) Did any special problems arise during the session (e.g., non-adherence to homework, interpersonal issues between therapist and patient, hopelessness about continuing therapy, relapse?)

YES NO

___ (b) If yes:

0 Therapist could not deal adequately with special problems that arose.

2 Therapist dealt with special problems adequately, but used strategies or conceptualizations inconsistent with cognitive therapy.

4 Therapist attempted to deal with special problems using a cognitive framework and was moderately skillful in applying techniques.

6 Therapist was very skillful at handling special problems using cognitive therapy framework.

13. Were there any significant unusual factors in this session that you feel justified the therapist's departure from the standard approach measured by this scale?

YES (Please explain below) NO

Part IV. OVERALL RATINGS AND COMMENTS

14. How would you rate the clinician overall in this session, as a cognitive therapist?

0	1	2	3	4	5	6
Poor	Barely Adequate	Mediocre	Satisfactory	Good	Very Good	Excellent

15. If you were conducting an outcome study in cognitive therapy, do you think you would select this therapist to participate at this time (assuming this session is typical?)

0	1	2	3	4
Definitely Not	Probably Not	Uncertain – Borderline	Probably Yes	Definitely Yes

16. How difficult did you feel this patient was to work with?

0	1	2	3	4	5	6
Not difficult -Very receptive			Moderately difficult			Extremely difficult

17. COMMENTS AND SUGGESTIONS FOR THERAPIST'S IMPROVEMENT:

18. OVERALL RATING:

Rating Scale:

0	1	2	3	4	5
Inadequate	Mediocre	Satisfactory	Good	Very Good	Excellent

Using the scale above, please give an overall rating of this therapist's skills as demonstrated on this tape. Please circle the appropriate number.

For instructions on the use of this scale, see: Young J.E., & Beck, A.T. (August, 1980). Cognitive Therapy Scale Rating Manual.

APPENDIX D

HOMEWORK RATING SCALE

HRS II

Instructions: This questionnaire consists of 12 questions regarding the client's homework completion from last session. Please read each question carefully, and circle the number of the **one response** that best describes your impression of the client's experience. If several statements apply equally well, circle the lowest number for that group. Be sure not to choose more than one response for any question.

1. Quantity
The client was able to do the activity
- 0 not at all
- 1 a little
- 2 some
- 3 a lot
- 4 completely

2. Quality
The client was able to do the activity well
- 0 not at all
- 1 somewhat
- 2 moderately
- 3 very
- 4 extremely

3. Difficulty
The activity was difficult for the client
- 0 not at all
- 1 somewhat
- 2 moderately
- 3 very
- 4 extremely

4. Obstacles
The client experienced obstacles in doing the activity
- 0 not at all
- 1 a little
- 2 some
- 3 a lot
- 4 extensive

5. Comprehension
The client understood what to do for the activity
- 0 not at all
- 1 a little
- 2 somewhat
- 3 a lot
- 4 completely

6. Rationale
The reason for doing the activity was clear to the client
- 0 not at all
- 1 somewhat
- 2 moderately
- 3 very
- 4 completely

7. Collaboration
The client had an active role in planning the activity
- 0 not at all
- 1 a little
- 2 some
- 3 a lot
- 4 extensive

8. Specificity
The guidelines for how to carry out the activity were specific
- 0 not at all
- 1 somewhat
- 2 moderately
- 3 very
- 4 extremely

9. Match with Therapy Goals
The activity matched with the client's goals for therapy
- 0 not at all
- 1 a little
- 2 somewhat
- 3 a lot
- 4 completely

10. Pleasure
The client enjoyed the activity
- 0 not at all
- 1 a little
- 2 somewhat
- 3 a lot
- 4 extremely

11. Mastery
The client gained a sense of control over their problems
- 0 not at all
- 1 a little
- 2 somewhat
- 3 a lot
- 4 extensively

12. Progress
The activity helped with the client's progress in therapy
- 0 not at all
- 1 a little
- 2 somewhat
- 3 a lot
- 4 extremely

APPENDIX E

ASSIGNMENT COMPLIANCE RATING SCALE

Assignment Compliance Rating Scale*

Directions: Using the following rating scale, determine the extent to which the client complied with the homework assignments given during the preceding session. Circle the number of the description corresponding to your judgment of the degree of performance. Please make an overall rating for all Homework Assignment(s) prescribed last session. Thank you.

1. The client did not attempt the assigned homework.
2. The client attempted the assigned homework but was unable to execute it for reasons such as lack of ability or extenuating circumstances.
3. The client did homework that was different from that assigned, but that would be considered 'relevant' to therapy and the client's particular target problems; e.g., the client had been given the assignment to challenge the idea that he or she is physically unattractive by writing a history of when friends and/or romantic partners have contradicted this idea. Instead, the client chose to ask the person he/ she is dating currently what they think of their attractiveness.
4. The client did a portion of the assigned homework.
5. The client did the assigned homework.
6. The client did more of the assigned homework than as requested.

APPENDIX F

MULTIPLE REGRESSION ANALYSIS HAACS ADHERENCE

Multiple Regression Analysis

To investigate which adherence behaviours of the HAACS best predicted overall score, stepwise multiple regression analysis was used to evaluate which individual adherence items would make greater contributions to overall adherence. The dependant variable, AI, was regressed upon the individual adherence items. Table 10 presents the means, standard deviations and intercorrelations beween the variables. As expected, adherence item ratings were positively correlated to overall adherence (AI). Included in the table are the values of partial correlations between each predictor and the dependant variable (i.e., HAACS AI), while controlling for all other predictors.

Multicollinearity was assessed based on an examination of the tolerance and variance inflation factor (VIF) statistics. Tolerance for HAACS items ranged from .23 to .76, and all VIF were greater than 1.40. These results suggest that the variables were not strongly associated with the other independent variables, thus satisfying an assumption of multiple regression analysis (Dancey & Reidy, 2002; Hair; Norusis, 2006).

A multiple regression analysis was conducted using the stepwise procedure in order to evaluate how well the adherence items predicted overall adherence. The predictors, or regressors, were the 18 HAACS adherence items (item 13 was removed from the analysis due to non-occurrence of behaviour), while the dependant variable was the overall adherence index (AI). The linear combination of adherence items was significantly related to the adherence index, $F(18,50) = 1388.63$, $p < .001$. The sample multiple R was .999, indicating that approximately 100% of the variance in AI

in the sample can be accounted for by the linear combination of adherence items. Together, the 18 adherence items accounted for 99.7 % of the variation in AI (adjusted R^2), indicating very good fit, however this is likely to reflect the inclusion of all the adherence predictor variables, rather than creating a parsimonious model (Hair). Due to the small sample size and large number of predictor variables, it was important to accurately test the assumption of independent errors (Norusis, 2006). The Durbin-Watson test was 2.24, which was within the range of 0.980 to 2.323 (p = .05) recommended by Savin and White (1977) when using a sample of this size with 18 regresssors. This indicates that the sample met the assumption of independence of errors.

An examination of the standardised Beta weights shows that item 3 (*use a situational conceptualisation to review previously assigned behaviour*) has the biggest impact on overall adherence in relation to the other independent variables in this model (β = .17). That is, for each rise of 1 standard deviation in item 3, overall adherence rises .17 SD. Item 4 (*use an individualised conceptualisation to make sense of any portion of non-completed homework*) provides the least impact on overall adherence (β = .06).

Table 25

Prediction of HAACS AI Scores Using Mean Adherence Item Ratings on the HAACS based on Multiple Regression Analysis

Variable[a]	Beta	SEB	β	*t**	95% CI of Beta	
					Lower	Upper
9	5.38	0.58	0.12	9.23	4.21	6.56
15	5.98	0.50	0.12	12.07	4.98	6.97
17	5.78	0.50	0.11	11.60	4.78	6.78
3	6.20	0.43	0.17	14.48	5.33	7.06
10	5.06	0.42	0.13	12.06	4.22	5.90
4	5.22	0.69	0.06	7.58	3.84	6.60
19	5.46	0.30	0.15	17.98	4.85	6.07
12	5.52	0.34	0.12	16.18	4.84	6.21
11	5.57	0.45	0.12	12.41	4.67	6.47
16	5.02	0.33	0.13	15.28	4.36	5.68
1	4.54	0.52	0.10	8.74	3.50	5.59
5	5.43	0.47	0.09	11.52	4.48	6.37
6	5.78	0.62	0.11	9.35	4.54	7.02
18	5.12	0.42	0.11	12.31	4.29	5.96
2	5.21	0.48	0.14	10.89	4.25	6.18
7	4.85	0.41	0.12	11.94	4.04	5.67
8	5.04	0.45	0.11	11.34	4.15	5.94
14	5.56	0.60	0.07	9.33	4.36	6.76

Note. n = 69. * All *t*-tests were significant, $p < .001$. Variables are listed in the order of entry using the stepwise method.

[a] Dependant variable is Adherence Index.

APPENDIX G

EQUATIONS FOR BLAND-ALTMAN PROCEDURE

Equations used for Bland-Altman Calculations.

Taken from Bland and Altman (1986, 2003) and Rankin and Stokes (1998)

Standard error of differences

$$\bar{d} = \sqrt{\frac{s^2}{n}}$$

Limits of agreement

$$\bar{d} \pm 2s$$

Bias between measures(95% CI $\bar{d}$, where t is t value for , n is number of subjects, and SE is

$$SE = \frac{SD_{diff}}{\sqrt{n}}$$

$$95\%\ CI\ \bar{d} = \bar{d} \pm t_{n-1} SE\ of\ \bar{d}$$

APPENDIX H

PEARSON'S *r* CORRELATION MATRIX FOR HRS-II SCORES

Table 26
Correlation Matrix of HRS-II items using Pearson's r

Items	1	2	3	4	5	6	7	8	9	10	11	12
1 Quantity												
2 Quality	.897(**)											
3 Difficulty	.494(**)	.550(**)										
4 Obstacles	.676(**)	.737(**)	.660(**)									
5 Comprehension	.694(**)	.764(**)	.502(**)	.530(**)								
6 Rational	.320(*)	.377(**)	.290(*)	.402(**)	.451(**)							
7 Collaboration	.007	-.133	-.023	.014	.077	.209						
8 Specificity	.088	.262	.130	.183	-.024	-.065	-.634(**)					
9 Match	.321(*)	.356(**)	.267(*)	.306(*)	.363(**)	.329(*)	-.012	.181				
10 Pleasure	.510(**)	.601(**)	.430(**)	.464(**)	.581(**)	.444(**)	-.097	.192	.452(**)			
11 Mastery	.583(**)	.706(**)	.391(**)	.558(**)	.518(**)	.465(**)	-.068	.208	.467(**)	.768(**)		
12 Progress	.675(**)	.786(**)	.488(**)	.596(**)	.594(**)	.410(**)	-.098	.287(*)	.521(**)	.775(**)	.923(**)	

Note. n = 63. Items listed are the items from the Homework Rating Scale. Correlations calculated usin Pearsons Product Moment
** Correlation is significant at the 0.01 level (2-tailed).
* Correlation is significant at the 0.05 level (2-tailed).

APPENDIX I

SCREE PLOT AND UNROTATED LOADINGS FOR HRS-II FACTOR ANALYSIS

Table 27
Component Matrix HRS-II(a)

	Component		
Item	1	2	3
quality	.909	-.061	-.258
progress	.904	-.113	.196
mastery	.849	-.060	.299
quantity	.821	.094	-.325
pleasure	.795	-.055	.315
obstacles	.780	.056	-.347
comprehension	.779	.242	-.147
difficulty	.656	.045	-.388
match with goals	.549	-.040	.514
rationale	.548	.372	.324
collaboration	-.066	.886	.068
specificity	.244	-.856	.019

Note. Extraction Method: Principal Component Analysis.
a 3 components extracted.

Table 28
Total Variance Explained by Eigenvalues

	Initial Eigenvalues			Rotation Sums of Squared Loadings		
Component	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	5.981	49.841	49.841	3.925	32.710	32.710
2	1.754	14.619	64.460	3.086	25.720	58.430
3	1.066	8.885	73.345	1.790	14.916	73.345
4	.708	5.901	79.247			
5	.627	5.224	84.470			
6	.547	4.556	89.026			
7	.483	4.024	93.050			
8	.305	2.538	95.588			
9	.232	1.929	97.517			
10	.194	1.614	99.131			
11	.054	.448	99.578			
12	.051	.422	100.000			

Note. Extraction Method: Principal Component Analysis.

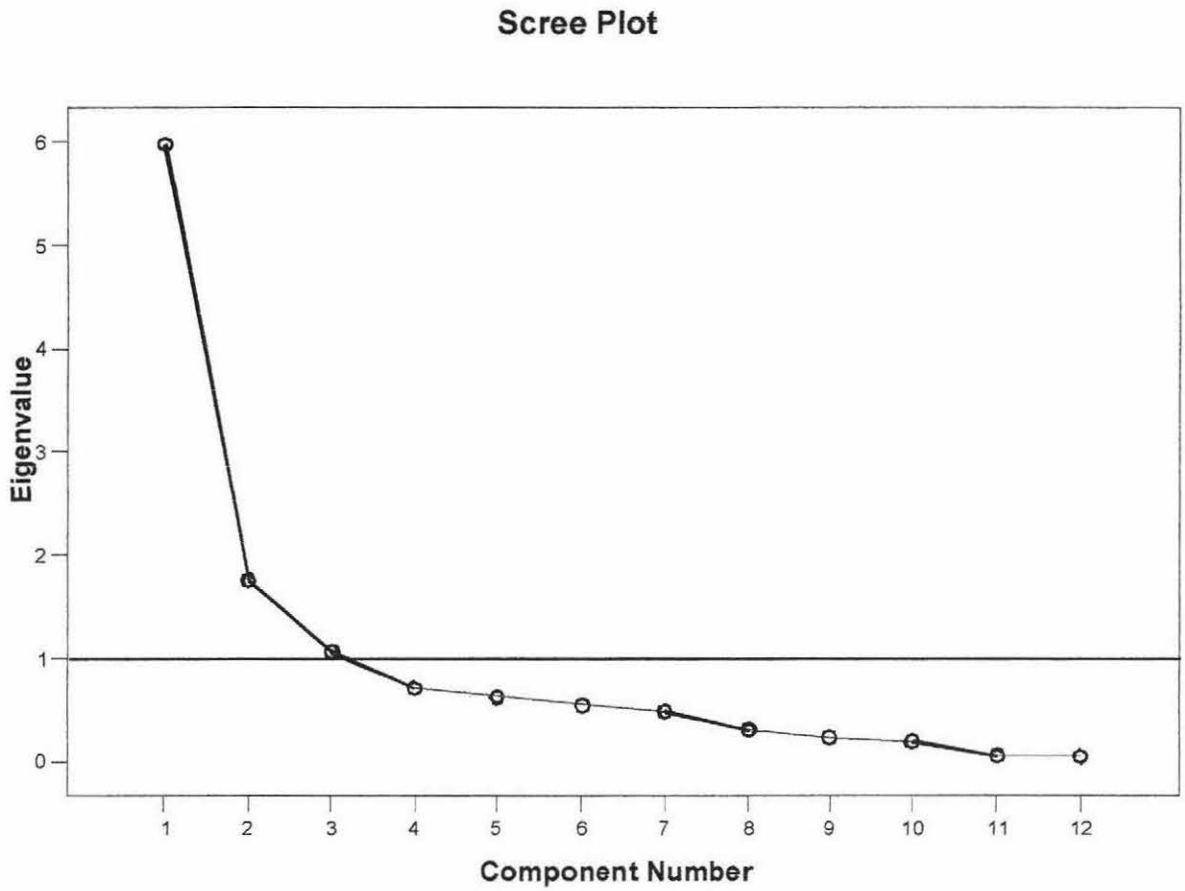

*Figure 10.*Scree plot of factor analysis for HRS-II.