

<http://researchcommons.waikato.ac.nz/>

## Research Commons at the University of Waikato

### Copyright Statement:

The digital copy of this thesis is protected by the Copyright Act 1994 (New Zealand).

The thesis may be consulted by you, provided you comply with the provisions of the Act and the following conditions of use:

- Any use you make of these documents or images must be for research or private study purposes only, and you may not make them available to any other person.
- Authors control the copyright of their thesis. You will recognise the author's right to be identified as the author of the thesis, and due acknowledgement will be made to the author where appropriate.
- You will obtain the author's permission before publishing any material from the thesis.

**Ultra-Brief Clinical Outcome Evaluations in New Zealand:  
A New Zealand Perspective on the Outcome Rating Scale and  
Session Rating Scale**

A thesis

Submitted in partial fulfilment  
of the requirements for the degree

of

**Master of Arts in Psychology**

at

**The University of Waikato**

by

**REBECCA PAY**



THE UNIVERSITY OF  
**WAIKATO**  
*Te Whare Wānanga o Waikato*

2017

## Abstract

Outcome evaluations are becoming a routine aspect of psychotherapy, yet many measures are deemed inappropriate for everyday clinical use.

Outcome evaluations help to inform clinical decision making and enhance treatment effects, whilst guiding therapy and tracking the client's progress.

Many current outcome evaluations are overly complex, lengthy and are not providing information in the most effective format possible. Given that mental health services nationally are increasingly being accessed, it is essential that the most effective and feasible outcome evaluations are being administered to maintain the meaningfulness of therapy.

The present study explored an ultra-brief treatment session outcome measure called the Outcome Rating Scale (ORS), and compared it to the current New Zealand mandated outcome evaluation Health of Our Nation Outcome Scale (HoNOSCA) ( $N = 98$ ). A quantitative, non-experimental correlational approach was taken, as studied in a clinical sample. Data was obtained through Lakes District Health Board, Infant, Child and Adolescent Mental Health Services.

The findings suggest that the ORS measures change and that it is comparable to HoNOSCA on 4 of the 13 HoNOSCA items. Interestingly, these 4 HoNOSCA items all significantly correlate with ORS's 'Overall' measure. Higher ORS scores positively correlate with higher SRS scores, highlighting the importance of therapeutic alliance. Maori intake scores are slightly lower than European intake scores, but even out following therapy. Implications for future clinical directions and future research are discussed.

## **Acknowledgments**

A special thanks to everyone who has supported me throughout this research process. Whether you helped me with concepts and ideas, or allowed me to relax and have a laugh – I cannot thank you enough.

Firstly, I would like to give my sincere thanks to the Lakes District Health Board, specifically George Furstenburg and Michael O'Connell for sharing their clinical data with me. I would also like to give my sincere thanks to all the participants. Without you this thesis would have been impossible.

Secondly, I would like to express my appreciation and thanks to my Academic Supervisor Dr Armon Tamatea, for his invaluable time and commitment in guiding me throughout the research process. You have taught me numerous insights on the process of undertaking academic research. I truly appreciate your guidance throughout the past year.

Thirdly, I would like to thank Waikato University's Psychology Research and Ethics Committee for approving this research and for ensuring the safety and wellbeing of all participants.

Finally, I would like to acknowledge my family and friends, for their continued support, encouragement, and advice throughout this research. The fun and laughs you have provided is what has kept me going. You are all one in a million. A special thanks to Simon, your proofreading skills and unconditional support was not taken for granted. Thank you for encouraging me.

## Table of Contents

Abstract.....	ii
Acknowledgments .....	iii
Table of Contents .....	iv
List of Tables .....	vi
Chapter One: Literature Review .....	1
Psychotherapy Defined .....	2
Theories of Psychotherapy .....	3
The Effectiveness of Psychotherapy .....	8
The Common Factors and Process of Psychotherapy .....	11
Client Variables as a Change Agent.....	13
Therapist Variables as a Change Agent.....	15
Routine Outcome Measurements.....	17
Characteristics of Outcome Measurements .....	19
Comprehensive Scales .....	23
Ultra-Brief Outcome Scales.....	23
Examples of Brief Outcome Measures .....	24
Outcome Questionnaire – 45.2 .....	24
Health of Our Nation Outcome Scale .....	26
Outcome Rating Scale .....	29

The Present Study .....	31
Chapter Two: Method .....	34
Setting.....	34
Participants .....	34
Data Screening.....	36
Measures .....	37
Health of Our Nation Outcome Scale .....	38
Outcome Rating Scale and Session Rating Scale .....	39
Research Design.....	42
Procedures .....	44
Data Analysis.....	45
Ethical Consideration .....	45
Ethics of Care .....	46
Relationships .....	49
Cultural Considerations .....	49
Chapter Three: Results .....	52
Descriptive Statistics .....	52
Ethnicity and the Outcome Rating Scale .....	54
Correlations .....	55
Hypothesis One.....	55
Hypothesis Two .....	57

Hypothesis Three .....	59
Feasibility of the Outcome Rating Scale .....	61
Chapter Four: Discussion .....	62
Hypothesis Testing .....	63
Hypothesis One .....	63
Hypothesis Two .....	64
Hypothesis Three .....	66
Feasibility .....	67
The Literature and the Present Study .....	68
Psychotherapy .....	68
Outcome Measures .....	71
Practical Implications .....	73
Study Limitations .....	75
Future Directions for Research .....	77
Final Conclusions .....	79
References .....	81
Appendices .....	92
Appendix A. Outcome Rating Scale .....	92
Appendix B. Session Rating Scale .....	93

## List of Tables

Table 1: Psychotherapy Theories.....	6
Table 2: Psychometric Property Descriptions.....	20
Table 3: Demographics of the Sample.....	44
Table 4: Comparison of ORS, HoNOSCA and SRS Pre- and Post- Data.....	53
Table 5: Gender Comparisons of ORS Means and Standard Deviations.....	53
Table 6: Maori and European Comparisons .....	55
Table 7: Comparisons Correlations Matrix Between Individual ORS Items .....	56
Table 8: Correlation Matrix for Pre- and Post- ORS and HoNOSCA Totals .....	57
Table 9: Session Rating Scales Means and Standard Deviations .....	60
Table 10: Correlation Matrix of SRS Scores and Pre- and Post- ORS Scores.....	60



## Chapter One: Literature Review

*“In truth... consumers (and payers) care little about how change comes about – they simply want it and in the most accessible format possible”*  
(Miller, Duncan, Brown, Sorrell & Chalk, 2006).

Psychotherapy has historically been criticised for being no more effective than not being treated (Eysenck, 1952). This has resulted in a growing interest in demonstrating psychotherapy's effectiveness and placed an emphasis on outcome evaluations becoming a routine system in mental health care services (Miller, Duncan, Brown, Sparks & Claud, 2003).

There is currently a range of methods for evaluating mental health care (Campbell & Hemsley, 2009). Many incorporate several elements in assessments, including both client and clinician ratings, observer ratings and teacher reports. However, numerous people in mental health care, such as Miller et al. (2006) suggest that although some of these outcome evaluation approaches are valid and reliable, criticisms regarding their motives, methodological complexity, therapeutic scope, administration length, and cost can make them infeasible in everyday clinical settings (Moran, 2017; Miller, 2012). These factors make outcome evaluations less likely to be used effectively and hinder therapeutic progress. These criticisms, along with arguments about the effectiveness of psychotherapy, highlight the need for not only valid and reliable routine outcome measures, but also outcome measures that are suitable for current clinical environments.

This study is about the use of an ultra-brief treatment session outcome measure, the Outcome Rating Scale (ORS; Miller & Duncan, 2000; see appendix) with a New Zealand mental health outpatient population. Further, this study aimed to explore how the ORS could fit into our current mental health services in New Zealand.

This chapter reviews the literature relevant to therapeutic outcome evaluations. It begins by defining psychotherapy, what it is, its point and how and why psychotherapy theories were refined or replaced overtime. The effectiveness of psychotherapy will be discussed with reference to the relationship between effectiveness and outcomes. Next, brief outcome measures will be presented in a discussion on their emergence and need. Finally, the present study will be introduced.

## **Psychotherapy Defined**

Outcome evaluations are used in the context of psychotherapy. There has been a mass increase of psychotherapies since the 1950's, each claiming effectiveness and bringing a new meaning to psychotherapy (Prochaska & Norcross, 2014). Due to the many different theoretical orientations a therapist can have, there is no single universal definition of psychotherapy (Wampold & Imel, 2015).

Defining psychotherapy depends on the therapist and the era the definition was developed in (Wampold & Imel, 2015). Arguably, psychotherapy has always been about assisting people, however, the route taken to achieve this has differed (Norcross, 2005). In 1949 at the Boulder Conference, psychologists were struggling how to define psychotherapy, Lehner concluded psychotherapy as, "we have left therapy

as an undefined technique which is applied to unspecified problems with a nonpredictable outcome. For this technique, we recommend rigorous training” (1952, p547). In comparison, a current working definition of psychotherapy by Norcross is, "Psychotherapy is the informed and intentional application of clinical methods and interpersonal stances derived from established psychological principles for the purpose of assisting people to modify their behaviours, cognitions, emotions, and/or other personal characteristics in directions that the participants deem desirable” (Campbell, Norcross, Vasquez & Kaslow, 2013, p.218).

These two definitions starkly contrast and highlight the development of psychotherapy in the past 70 years. Both definitions suggest that psychotherapies purpose is to assist people. However, the historic definition provides little guidance and leaves a lot to be interpreted. It reflects early thinking of psychotherapy when therapies were not standardised, therefore no specific techniques or outcomes could be explained. The current definition is broad, balanced and neutral in regard to theory and method but emphasises that psychotherapy is based on the application of informed clinical methods and psychological principles. It incorporates direction on how and why psychotherapy will be effective and specifies areas of change. The areas of change in psychotherapy are what is measured in therapeutic outcome evaluations. This will be discussed in further depth later in the chapter.

### **Theories of Psychotherapy**

Selecting a specific psychotherapy theory helps guide therapists to better understand their clients and their clients’ challenges, and it also

helps provide therapeutic solutions from a specific viewpoint (Lambert, 2004). Essentially, selecting a psychotherapy theory is about taking into account the theory of change that determines the selection of treatment modality. Without a theory, clinicians are directionless and vulnerable to ineffective treatment (Prochaska & Norcross, 2014).

Table 1 refers to different psychotherapy theories, and shows that as time has progressed new psychotherapy movements, like Cognitive Behaviour Therapy (CBT), have developed and they tend to become centred on holistic, integrative approaches (Wampold & Imel, 2015).

Historically, psychotherapy emerged in the context of modern medicine (Caplan, 1998). Early medicine disregarded mental health disorders and instead placed an emphasis on physiochemical (somatic) processes and left the cause as unknown (Taylor, 1999). In the mid to end of the 19<sup>th</sup> century psychotherapy was centred on mind cures (Caplan, 1998). This was mainly accessed through different religions, specifically, Christian Science. The idea was that physiochemical problems could be healed through the mind, faith or spirituality (Taylor, 1999). Largely, the psyche was ignored by health professionals till the field of medicine claimed in the interest of the public's safety, the psyche needed to be treated by a Doctor. By 1909 Freud's Psychoanalysis was seen as more scientific than mind cures and had started gaining popularity (Wampold & Imel, 2015).

Freud's Psychoanalysis dominated the field of psychotherapy till around the 1970's (Wilson, 1982). However, in the 1950's alternative psychotherapy models were being introduced as societies needs were

changing following the Second World War and a wave of research being produced (Wilson, 1982). Behaviour Therapy (BT) became popular in the 1970's as it rejected unobservable phenomena like the unconscious, defence mechanisms, intrapsychic structures and spirituality in psychoanalysis/psychodynamic theories (Eysenck, 2013). Thus, it provided an objective method of thinking about a client's presenting challenges. Following BT, a focus was placed on cognitive theories and eventually humanism theories emerged (Norcross, 2005). Humanism theories, such as Carl Roger's Person Centred Therapy, were an attempt to help people find meaning in life following the Second World War and Holocaust (Wampold & Imel, 2015). These theories all still have their strengths and weaknesses and many are still practiced especially under psychotherapies theory integration movement.

Integrative psychotherapy theories can be dated back to 1919 when Freud developed psychoanalytic theories as an alternative to classical analysis (Beit-Hallahmi, 1974). However, it has only been a specific area for the past 25 years (Norcross, 2005). Integrating psychotherapies is motivated by the want to learn and understand other theories, to see how clients can benefit, and to enhance the effectiveness and efficiency of treatment (Norcross, 2011). Integrative theories acknowledge the limitations of theories and possible benefits of others (Wampold & Imel, 2015). Integrative therapies have refined/replaced one school approaches over time for numerous reasons, such as, the mass increase of theories and the empirical research behind them, an understanding that one theory does not suit all clients or challenges, a change in the type of therapy used

(short-term, problem-focussed treatments) and a recognition that common features (client factors, relationship factors, hope and expectancy) contribute to treatment efficacy (Norcross, 2005).

Table 1

*Contemporary Psychotherapy Theories and their Essential Qualities*

Name of Theory	Founder	Core Assumptions	Strengths	Limitations
Psychoanalysis Theories	Sigmund Freud (Late 1800's)	A focus on unconscious meanings & motives of behaviours and thoughts. The goal is to change problematic behaviours or thoughts through better understanding their unconscious drives.	A comprehensive theory that withstood the test of time. Revolutionary in its time.	Does not really use common sense, is now dated and does not quite fit with our current knowledge, unfalsifiable.
Psychodynamic Theories	Sigmund Freud (1900's)	Similar to Psychoanalysis, but a greater focus on present behaviours and changes.	Highlighted the importance of childhood, defence mechanisms, made the case study method popular in psychology	Unscientific, too deterministic, rejects free will, unfalsifiable.
Behaviour Theories	Ivan Pavlov, B. F. Skinner & Watson (1913)	Behaviours are learned, therefore they can be unlearned. A focus on the present and not interested in unconscious motives. Therapists replace the maladaptive	Scientific, very applicable, emphasise objective measurement.	Does not acknowledge mediational processes, too deterministic, humanism – comparing animals to humans.

		behaviour with more effective behaviour strategies.		
Cognitive Theories	Aaron T. Beck (1960s)	Look at how a person's conscious thoughts contribute to their challenges. It is believed that problematic thinking causes problematic functioning in behaviours/emotions. If you can change a person's dysfunctional thought process, then you can change how they feel and behave.	It highlights people's thought processes, can be combined easily with other approaches.	Does not acknowledge factors of basic human behaviours and is based on controlled experiments.
Cognitive Behaviour Theories	Aaron T. Beck & Albert Ellis (1950/1960s)	Connecting behaviour therapy and cognitive therapy together. The idea that thoughts impact behaviour and targeting both in therapy.	It focusses on human thought, has been proven to be effective.	The exact role of cognitive processes is not quite known, can be seen as narrow thinking as it does not include all areas of human functioning.
Humanistic Theories	Carl Rogers, Abraham Maslow & Otto Rank (1960s)	Believe in freedom of choice. That people are capable of making rational choices to reach their full potential. Through taking responsibility for yourself you can find your meaning and process through challenges.	It emphasises individual choice and responsibility, considers the person in their environment with their perceptions and feelings.	It is considered a vague theory, too subjective to verify often considered not to be 'true' science

Integrative Theories	Wachtel's & Ryle's, among many others (1990's -)	Combining different positive elements of several theories and using them together in therapy. By individualising therapy and choosing techniques from different perspectives that best suit their client, therapy is said to be more meaningful.	Individualised therapy, using current scientific findings, targeting individual needs.	Can become overly complicated, based on what therapist sees as 'right'
----------------------	--	--	--	--

### The Effectiveness of Psychotherapy

Early efforts to establish treatment efficacy began in the 1920's (Strupp & Howard, 1992). The lack of objective verification of outcomes caused people to doubt the efficacy of psychotherapy, and triggered movement in efficacy studies (Wampold & Imel, 2015). The first direct observation of treatment was in the 1940s by Carl Rogers and colleagues (Rogers, 1951). They transcribed sessions from audio tapes and developed hypotheses regarding whether psychotherapy resulted in changes in personality (Rogers, 1951). However, their sample population was small, treatment was obscure, and disorders and outcomes were not well defined (Goldfried & Wolf, 1996). This complicated research on psychotherapy and clouded any findings on its effectiveness. Following Rogers research, Eysenck claimed that the rate of recovery of clients accessing psychotherapy was equivalent to the rate of random remission. Suggesting psychotherapy is not effective, but purely based on chance (Eysenck, 1952). This comment intensified the debate of psychotherapies



effectiveness and its research methods (Ludy, 2007). Research methods such as randomised design, placebo control group design, and meta-analysis were created to support scientific methods to provide evidence of psychotherapies efficacy.

Meta-Analysis is a research method that combines and averages results across studies on specific topics (Howard, Moras, Brill, Martinovich & Lutz, 1996). The first meta-analysis on psychotherapy was by Smith Glass, and Miller (1980). It analysed all studies that compared a psychotherapeutic approach to some type of control group. It found psychotherapy to be effective with a mean effect size of 0.85 (Smith et al. 1980).

The meta-analysis by Smith et al. (1980) reviewed 475 waiting list or placebo group controlled studies of various diagnoses and treatments with satisfactory quality. They reviewed numerous types of psychotherapy (Psychoanalysis, BT, CBT, for example) compared to untreated client groups and found most to be effective, especially CBT, cognitive therapy and hypnotherapy. Smith et al. (1980) had an average effect size of 0.85 immediately after treatment and it went down to 0.5 on average at a two-year follow-up. Suggesting, overall psychotherapy is most effective immediately after treatment and it loses some effectiveness once therapy finishes. This meta-analysis also found that fear and anxiety were generally the most responsive to treatment, and measures of personality traits and work or school performance were the least responsive. It indicated that people who took antipsychotic medication experienced the least improvement and clients with depression and clients with simple

phobias experienced the most improvement. Smith et al. (1980) also found females benefitted more from treatment than men. Overall, Smith et al. (1980) found that at the end of treatment the average client is 80% better than the untreated client. Psychotherapy is still regarded as effective across numerous disorders and therapies. More recently, Lambert (2013) revisited Eysenck's (1952) study to evaluate the efficacy findings on psychotherapy and specific treatments, such as CBT. Lambert (2013) found that the average effect size between treated and untreated populations is  $d = .75$ , which is a general success rate of 67%, compared to 33% for an untreated population.

The evidence supporting psychotherapies effectiveness meant research no longer needed to show the efficacy of psychotherapy in general. Rather, research has turned to finding evidence to support specific treatments and their efficacy with specific disorders (Wampold & Imel, 2015). Outcome evaluations were brought in to track change, and look at how each individual client is functioning, and what treatment will best target their disorder/challenge(s). It was through this thinking where standardised treatments and manuals were created to provide guidelines to reduce variability and increase the effectiveness of psychotherapy. This movement can also account for psychology's focus in the 1990's on using empirically supported treatments (ESTs). ESTs were treatments that had evidence supporting their effectiveness (Barlow, 2004). It was believed that psychotherapy was only effective if one of the ESTs were used. More recently, psychotherapy has based effective treatments on evidence-based practice, which refers to the integration of empirical research into

clinical practice, commonly known as the Scientist-Practitioner Model (Wampold, Goodheart, & Levant, 2007). In the Scientist-Practitioner Model the best research is used to inform treatment and the changes are monitored through outcome scales, and are used to help guide future research questions (Shapiro, 2002).

### **The Common Factors and Process of Psychotherapy**

Psychology has a history of rivalry amongst theoretical orientations. Historically, therapists only practiced from their theoretical framework and each theorist believed their theory was the best treatment (Rosenzweig, 1936). However, due to therapy outcomes being similar across therapies, evidence supported there are factors that occur in the process of all psychotherapy, called common factors (Wampold et al. 2001, 2007, 2010; Rosenzweig, 1936). Common factors include, client expectation, therapist relationship, hope and corrective experience. Common factors are considered a component of change and are a factor in the efficacy of psychotherapy (Lambert, 2004).

Effectiveness in psychotherapy can be considered as a positive change in awareness, behaviour and personality (Prochaska, 1982). Processes of change refer to the actions of people to change their cognitions and behaviours to address presenting challenges (Prochaska & Norcross, 2014). The processes of change can be thought of as strategies of psychotherapy theories that occur at an individual or an environmental level. The process of change that occurs during and between therapy is what is measured as effective therapy. The main processes of change are:

consciousness raising, catharsis, choosing, conditional stimuli and contingency control.

Consciousness raising helps clients be more aware of themselves and their environment (Wampold & Imel, 2015). When change occurs through increasing consciousness, more information is available to the individual, and they can choose the most effective responses. Similarly, in catharsis, change occurs through expressing emotions, and being able to understand your challenges and make corrections where needed (Prochaska & Norcross, 2014). Catharsis provides an insight into an individual's experience and helps clients to process emotional blocks. The power of choice can influence change and is possible through consciousness raising, and becoming aware of new alternatives to dealing with a challenging problem. Change can also occur through conditional stimuli. Conditional stimuli refer to modifying environmental or behavioural factors that control our responses (Prochaska & Norcross, 2014). Such as learning to do the healthy opposite of the behaviour (counterconditioning), or removing or avoiding social triggers that cause problem behaviours (stimulus control). These strategies refer to the processes in therapy that aid change and provide some explanation to what happens during effective therapy. For example, when a common factor such as therapeutic alliance is present in therapy, changes through consciousness raising may be more effective due to a client's trust in the therapist.

An important point emphasised by Janse, Boezen-Hilberdink, Van Dijk, Verbraak, and Hutschemaekers (2014) and Flannery-Schroeder and Lamb (2009) is that there is no one ideal therapy method or technique that

will continuously get an effective outcome, rather it differs on a case by case basis. This reflects the trend towards researching specific treatments for specific psychological disorders as well as using integrative therapies as outlined in Table 1. Lambert (2004) states that not all clients will make meaningful change, and some clients may not maintain their therapeutic progress. He suggested that this is due to not all clients having positive experiences with common factors (such as expectation for improvement, persuasion, warmth and attention, understanding and encouragement). Common factors, such as the ones above, in Lambert's opinion, are necessary for the processes of change to occur.

### **Client Variables as a Change Agent.**

Psychotherapy efficacy and outcomes are impacted by common factors in treatment and the process of change, however, there is a myriad of client variables that also impact on psychotherapy and its efficacy. All clients present with their own set of challenges ranging in severity, education, achievements, and willingness to change (Ardito & Rabellino, 2011). These variables impact on how the client functions, how they will engage in therapy and their therapeutic outcome.

### ***Diagnosis as the prescriptive client variable.***

Research has indicated that specific disorder characteristics are linked to therapeutic outcome success (Ardito & Rabellino, 2011). Ideally, all therapy should target the client's challenge(s) (diagnosis). The Diagnostic and Statistical Manual of Mental Disorders (DSM) was developed to provide systematic evaluations of mental health disorders,

and to provide guidance on what the disorder is as well as its characteristics (American Psychiatric Association, 2013). Although the DSM is not always viewed favourably due to labelling, this information can be used to provide the best course of action for treatment by being able to treat a specific diagnosis rather than a broad array of presenting challenges (Wampold, 2010). Wampold (2010) reports that treating specific disorders provides better chances of positive outcomes because it gives clinicians clarity and can help clients better understand and manage their challenges.

### ***Personality Variables.***

There are specific personality characteristics that are said to make treatment process and outcome more likely to be effective. These are expectancies (of positive outcomes), a readiness to change and ego strength. Cohen, Beard, and Björgvinsson (2015) found that a positive expectancy trait provided a better outcome in therapy. A part of this is the confidence a client has in their therapist. If a client trusts their therapist, they are more likely to expect treatment to work and their outcome will likely be more effective (Flannery-Schroeder, & Lamb, 2009). However, client expectancies have been strongly connected to treatment duration and less connected to treatment outcomes (Lambert, 2013). Suggesting, these positive expectancies develop over time as the therapeutic alliance develops.

A readiness to change is another personality factor that has a strong relationship with positive treatment outcomes (Flannery-Schroeder, & Lamb, 2009). Someone who wants to change and is trusting in their

treatment is more likely to change. In literature, this has been connected to ego strength. Ego strength refers to positive personality characteristics that allow a person to overcome their challenges, and learn new skills to better their functioning; somewhat of a resilience (Bohart & Wade, 2012).

### ***Interpersonal variables.***

Interpersonal relationships often refer to therapeutic alliance and are an important part of therapy and its outcomes (Flannery-Schroeder, & Lamb, 2009). Being able to build a healthy therapeutic alliance causes therapists to be trusted and facilitates discussion, resulting in treatment efficacy (Flannery-Schroeder, & Lamb, 2009). Bordin (1979) states that therapeutic alliance has three aspects: an agreement on goals, assignment of tasks, and the development of a bond. This emphasises that therapy is a continuous interpersonal process that requires positive therapeutic relationships for positive outcomes. Having outcome evaluations that speak to these variables would mean clients are more likely to be understood, treatment can cater to their specific needs, and change processes can occur.

### **Therapist Variables as a Change Agent.**

#### ***Observable States.***

Observable states refer to therapist training, experience and skill. Research has found that professional training, amount of training and skill has a positive impact on therapy outcomes (Flannery-Schroeder, & Lamb, 2009). However, professional experience has mixed research outcomes. Several researchers have found that experienced therapists had more

positive outcomes than inexperienced therapists (see Propst et al. 1994; Blatt et al. 1996), although the amount of time spent training had less effect than the amount of clinical time. Suggesting therapists with more clinical experience have better outcomes than therapists with mostly training experience. This may suggest therapists learn how to create common factors through clinical experience rather than in training. In addition, a therapist whom a client perceives as competent is more likely to get positive outcomes (Lambert, 2004). This may indicate that therapists who are competent in their treatment create an environment where clients feel safe and secure, supporting the processes of change and ultimately supporting treatment outcomes.

### ***Inferred Traits.***

Inferred traits refer to personality and coping mechanisms, therapists' emotional well-being, values, beliefs and cultural attitudes (Blow, 2008). Variables in inferred traits are a means of finding out how the therapist and client will interact. Psychotherapy is based on change, and throughout history, it has been criticised for the potential of therapists to place their own values, beliefs and attitudes onto their clients throughout the process of change (Norcross, 2011). However, results from research on values and beliefs of therapists are mixed with some showing clients who chose to find positives in the traits of their therapist linked to positive treatment outcomes, and others finding nothing significant (Norcross, 2011). Largely, it appears specific therapist traits are an individual preference and are not a global indicator of outcomes.



### ***Inferred States.***

Inferred states refer to the therapeutic relationship and are consistently connected to therapy outcomes (Lambert, 2004). The therapeutic relationship is the most consistent and widely accepted measure of therapeutic change and outcome (Lambert, 1992; Norcross, 2011). Lambert (2004) states that patient-therapist relationships account for a significant but relatively small percentage of treatment outcomes (on average 7% to 17%). Suggesting, psychotherapy is more than a warm relationship between client and therapist as suggested by Eysenck (1952) study, but rather an interpersonal process that relies on other processes of change, such as consciousness raising. Some outcome measures, such as the Session Rating Scale (SRS), have been designed to facilitate discussion about the therapeutic relationship. By discussing the relationship, changes can be made for the client to receive the most beneficial treatment. The SRS will be discussed in further depth later in the chapter.

### **Routine Outcome Measurement**

Routine outcome measurements (ROM) were introduced to mental health clinics in the early 1990's and since have been adopted by mainstream health providers as a necessary and required component of mental health services (Trauer, Gill, Pedwell & Slattery, 2006). ROMs commonly fall under a service goal of measuring the quality of care and reviewing outcomes of mental health and addiction services (Trauer et al. 2006). All clinicians and professionals who are to use the measurements are trained in the tool as well as given an understanding of why ROM is

necessary (Trauer et al. 2006). ROM is considered a necessary part of mental health services because of the benefits to all those involved in the service – users, carers, clinicians, health departments, government. It provided a platform to communicate perspectives, for clients to be readily involved in their care, for progress to be tracked, and ultimately, it provided more effective treatment outcomes (Callaly & Halleborne, 2001). For example, based on findings from Lambert (2013), using outcome evaluations reduced expected deterioration rates by 50% (to 9%) and increased positive therapy outcomes from 22% to 38%, and improved overall treatment results.

However, even with effective treatment outcomes, ROM has historically been criticised and many clinicians still struggle to see why it is important and a required part of mental health services (Trauer et al. 2006). Some clinicians believe ROM is not actually used for its intended purpose of aiding therapy but rather interferes with the therapeutic alliance (Boswell, White, Sims, Harrist & Romans, 2013). This is supported by staff attitudes in a British clinic where it was said that ROM provided a too simplistic view on care, results were not always used for treatment and planning, and ROM detracts from the therapeutic relationship (Cuijpers, Li, Hofmann & Andersson, 2010). Although they mentioned that if ROMs have good reliability and feasibility they were more likely to be used. This suggests a need for highly feasible ROMs to get clinicians engaging with them and reaping their benefits. It is stated across ROM literature that as time passes and new staff are introduced, ROM will be perceived as ‘just a part of the system’ (Boswell et al. 2013).

A large body of research indicates that the client's subjective experience of change early in the treatment process is one of the better predictors of treatment outcome (Littauer, Sexton & Wynn, 2005; Lambert & Bergin, 1994). Lambert et al. (2001) split 609 clients into four groups, either experimental or control groups to measure how feedback about client progress impacted client progress and their session attendance rates. Lambert and colleagues (2001) found that incorporating outcome information into therapy resulted in a 65% improvement in the success (an increase in client outcome scores) of cases most at risk of a negative outcome. In another study of over 3,000 cases at a single agency, the ongoing use of outcome information over 1 calendar year resulted in a 150% improvement in overall effectiveness (Miller et al. 2003). These findings suggest that outcome evaluations are necessary in clinics to ensure change is occurring and to give treatment the best chance of succeeding.

### **Characteristics of Outcome Measures**

It is largely agreed that at a minimum, all outcome evaluations require good psychometric properties, this includes reliability and validity, but also, being sensitive to change, and feasible in clinical settings (Miller & Duncan, 2000). Psychometric properties refer to the usefulness and appropriateness of the measures, and are commonly split into four areas (reliability, validity, sensitivity to change and feasibility) (Coaley, 2009). Reliability refers to the consistency of the measure (Te Pou, 2012). For example, how sure we are that the measure will consistently provide the same result over time and in any situation. Some types of reliability are,

internal consistency, inter-rater reliability and test-retest reliability. Validity refers to whether the tool actually measures what it is supposed to (Coaley, 2009). For example, validity looks at what the test measures (such as risk) and how well it measures the construct. There are three main areas of validity, these are, construct, content and criterion. Sensitivity to change refers to whether the measure reflects change over time, and feasibility refers to the degree the measure is used and accepted in clinical environments (Miller et al. 2003). Reliability and validity are intertwining concepts. Meaning, for a measure to be considered appropriate, it needs to be both valid and reliable, for example it must reliably measure the construct to be valid. Table 2 displays some specific psychometric properties.

Table 2

*Psychometric Property Descriptions.*

Psychometric Properties	Description
Content validity	The degree the measure reflects its area of interest. For example, the items its measuring should reflect traits of the overall construct/topic it intends to measure.
Criterion validity	Shows the accuracy of a measure by comparing it to another tool that measures the same construct. The two types of criterion validity are concurrent – when you compare a measure to an existing measure with known sound psychometric

---

	properties, and predictive validity – a measures capability of measuring a future outcome.
Construct Validity	Refers to how well the measures accurately measures its construct.
Internal Consistency	Refers to the similarity of items across a measure.
Inter-rater Reliability	Looks at the consistency across administrators scores. High inter-rater reliability suggests the measure provides a valid measure of the construct.
Test-retest Reliability	A measure of an individual's score consistency across two administrations of the measure. It indicates how reliable the measure is over time.

---

There are numerous types of measures and how different measures collect information can impact on the validity, reliability and feasibility of outcome scales. Outcome measures can be comprehensive or brief and completed by the client, therapist or by another person who knows the client, for instance a parent or teacher (Coaley, 2009).

Self-report outcome measures refer to any outcome evaluation where clients complete the scale by themselves without any input from a therapist or another person (Lambert, 2004). Self-report outcome scales are a good way to quantify symptomatology and to examine change over time. They allow the client to put their state, behaviours and feelings into

their own words and give them some control over their treatment and care (Miller et al. 2006).

Clinician-rated outcome scales refer to outcome scales that are rated independently through the therapist (Miller, 2012). In clinician-rated scales, clients do not physically see the scale as it is completed after the therapy session. The clinician uses the information gathered in the therapy session combined with their clinical judgement to complete the outcome evaluation. This information is then recorded to guide treatment and treatment planning, it is also used to measure therapeutic change (Cuijpers et al. 2010).

A challenge encountered in psychotherapy is that measurement can be driven by the values and biases of those providing the data (Boswell et al. 2013). It is suggested that clients' opinion on their assessment and treatment can differ vastly from the clinicians (Hafkenscheid, Duncan & Millers, 2010; Janse et al. 2014). This may be because clinicians have formulated incorrectly, clients may forget details of behaviours, attitudes, and symptoms or clients may be controlled by cognitive biases. Other people may lack the self-awareness required to answer some of the questions or have an ulterior motive. These factors impact outcome information and can lead to less effective therapy. Although accessing reliable information can be challenging, the main barrier to outcome evaluations is their feasibility.

### **Comprehensive Outcome Scales.**

Comprehensive outcome scales refer to scales that provide a detailed overview of a person's functioning (Campbell & Hemsley, 2009). Comprehensive outcome scales are known for excellent psychometric properties (good validity and reliability) but also for poor feasibility. The feasibility of comprehensive scales is often based on the length of time needed to administer, interpret and use in clinical settings (Coaley, 2009). Although comprehensive scales are not commonly viewed as user-friendly or practical, they do provide a depth of knowledge that outweighs the depth of information any other scale can obtain (Miller et al. 2003).

### **Ultra-Brief Outcome Scales.**

Ultra-brief outcome scales refer to scales that provide a simple overview of a person's functioning (Miller & Duncan, 2000). They are used routinely in therapy and broadly measure main known areas of wellbeing and they can be used immediately during clinical sessions (Miller et al. 2003). It is reported they are user-friendly, clear, concise and quick and easy to administer and interpret in therapy (Miller & Duncan, 2000). Brief scales are also usually low in cost making them practical for repeated use. Due to brief scales simple methodology, minimal training is usually required in the administration, scoring and interpreting of results (Campbell & Hemsley, 2009). Miller et al. (2006) stated ultra-brief outcome scales have adequate psychometric properties, and are highly feasible in clinics. However, ultra-brief outcome scales do not have the capacity to achieve the same accuracy and depth of information that a more complex scale does (Miller et al. 2003). A limitation of brief scales is they are not

designed to measure risk and will still need to be used alongside other measures to gain a full picture (Miller et al. 2003).

Although ultra-brief measures cannot access the same level of information as a comprehensive scale, they can have adequate psychometric properties and exceed the industry standard on what is acceptable for an outcome tool due to their high feasibility. Brief measures have emerged as mental health services are increasingly being accessed, meaning stricter time and money constraints (Boswell et al. 2013). Ultra-brief measures provide a cheap outcome tool that promise to reduce time spent doing paper work and increase the time spent in therapy (Miller, 2012). To make accurate clinical decisions and for treatment to be more effective, there needs to be easier access to scales that measure similar characteristics to complex measures, and compare to their validity and reliability, but in shorter more concise formats. The goal is by increasing the feasibility of outcome scales, the effectiveness of therapy will also increase.

### **Examples of Brief Measures**

#### **Outcome Questionnaire – 45.2.**

The Outcome Questionnaire – 45.2 (OQ-45.2) is a highly-regarded outcome evaluation (Boswell et al. 2013). It uses a 45-item self-report scale that is designed to measure functioning for repeated use throughout a client's course of therapy (Bringhurst, Watson, Miller & Duncan, 2006). The OQ-45.2 is usually administered online but can be completed by paper. The 45 items are measured on a five-point scale from “almost always” to “never” (Lambert, 2012). It takes roughly 5 to 10 minutes of



therapy time to administer and in theory, was created to be able to glance at and see what changes in mental health functioning are of importance for that session (Boswell et al. 2013). There is an extensive online system (OQ – Analyst) that scores and records the results that can be accessed in a variety of ways (Lambert, 2012). This software can have the functioning results (from all sessions), therapy suggestions and suggested outcomes on the screen within one second of completing the last item (Lambert, 2012).

The OQ-45.2 has sound psychometric properties which are the main reason for its popularity in clinical settings (Lambert, 2012). Bringhurst et al. (2006) reported the OQ-45.2 has high internal consistency (.93) and test-retest reliability (.84) when tested on a clinical, community and undergraduate American samples. Lambert, Burlingame, Umphress, Hansen, Vermeersch, Clouse, and Yanchar (1996) found that the OQ-45.2 is highly effective in accounting for the effects of therapy, due to it being sensitive to change in a treated population, and remaining stable in a non-treated population. Lambert et al. (1996) also report that the OQ-45.2 has moderate to high validity coefficients between other well-established outcome measures, such as Beck Depression Inventory, Taylor Manifest Anxiety Scale, and the SCL-90-R. Most recently, Boswell et al. (2013) tested the OQ-45.2 on 220 counselling centre clients at an American university. They reported the OQ-45.2 had acceptable psychometric properties, with high correlations for the total score and Symptom Distress scales, but moderate correlations Interpersonal Relations and Social Role subscales (Boswell et al. 2013). Compared to

the HoNOS, the OQ-45.2 has more positive effective outcome research results and appears to be more valid in routine clinical settings.

As previously mentioned, the OQ-45.2 is widely used and respected, however, it still has its challenges in clinical settings. Miller and Duncan (2004) reported that clinicians from busy clinics found the feasibility of the OQ-45.2 a limitation of its clinical use. They found the OQ-45.2 administration lengthy and at times too complex (Miller & Duncan, 2004). This resulted in it not being completed or putting other client's therapy sessions behind time and causing a general havoc in the clinic. Although the OQ-45.2 has sound psychometric properties, it still needs to be feasible in clinical settings for it to be administered. Longer outcome evaluations make this harder to achieve (due to the time constraints in clinics) which is why there is a growing need for ultra-brief outcome scales with similar reliability and validity measures.

### **The Health of the Nation Outcome Scales.**

The HoNOS is a tool used to measure mental health outcomes. It was designed by the United Kingdom Royal College of Psychiatrists Research Unit for routine clinical use (Brooks, 2000). The HoNOS measures are also used to collect information for New Zealand's mental health database, which collects information for the government to inform health funding (Te Pou, 2012). The HoNOS is the current Ministry of Health appointed outcome scale used in New Zealand (Brooks, 2000; Te Pou, 2012). This means that anyone who accesses mental health care in New Zealand will complete the scale throughout the course of their therapy. The HoNOS is used to measure the health and social functioning

of people accessing mental health services (Te Pou, 2014). There are three versions of HoNOS, HoNOSCA covers children and adolescents, HoNOS is used on working aged adults, and HoNOS65+ for people aged 65 years and older (Kisely, Campbell, Cartwright, Cox & Campbell, 2010). To administer HoNOS you need to be trained in the scale and the accompanying software (Brooks, 2000).

There are 12 items in HoNOS for adults, where HoNOSCA for children and adolescents has 13 items (Kisely, et al. 2010). All scales are independent measures; however, they are further broken down into 4 sections (behaviours, impairments, symptoms and social functioning). All scales are scored from 0 to 4, a score above 2 is clinically significant (Muller, Vandeleur, Weniger, Prinz, Vetter & Egger, 2016). The higher someone scores the higher the morbidity (Kisely, et al. 2010). HoNOS is designed to take 5 minutes to complete following a clinical session, although, some clinicians have reported otherwise (Brooks, 2000).

HoNOS has been questioned for its usefulness and whether it is suitable for routine clinical settings (Te Pou, 2012). Interestingly, Kisely, et al. (2010), highlight that HoNOS was designed to measure outcomes in clinical settings, although many studies on HoNOS have focussed on the instrument's cross-sectional features (such as reliability, criterion and concurrent validity), rather than assessing its sensitivity response to change over time. Kisely et al. (2010) developed a study using 4620 administrations of the HoNOS in outpatient mental health care data in Canada, to test the HoNOS sensitivity to change. Overall, they concluded that it is clinician friendly, sensitive to change and reliable. They found that

it was sensitive to change in 9 out of 12 adult version items in outpatient settings. Kisely and colleagues (2010) had an 82.2% completion rate which fell to 49% at follow-up. They found that improvement over time was most evident for the total score but was also highly visible in all other areas, particularly in the adult version, indicating sensitivity to change. They also found good predictive validity, suggesting HoNOS can provide a good indication of length of stay, readmission rates and retention in the community. Kisely and colleagues (2010) study suggest it is appropriate for routine use in clinical settings, however, Muller et al. (2016) suggest that HoNOS does not perform adequately in more severe and debilitating disorders. This limitation limits the reliability of the measure and questions its validity.

Te Pou (2012) qualitatively summarised peer-reviewed findings on 20 HoNOS articles that employed the same format as the one used in New Zealand. They reported that overall, HoNOS was easy to use, low cost, applicable to a wide range of mental health disorders and that it had some strong psychometrics. In particular, they found the HoNOS measure performs well when compared to other clinician administered measures, but performs poorly when compared to self-administered measures, such as the Symptom Checklist 90 Revised (SCL90-R) (Te Pou, 2012).

However, research conducted in Australia by Brooks (2000), who compared the HoNOS to the SCL90-R, found that HoNOS's validity is questionable, sighting that it did not correlate with "major measures of mental health symptoms or with measures of mental health status". Muller et al. (2016) also raised questions about HoNOS's validity, stating its

validity shifted depending on the diagnostic group. These findings are a cause for concern and highlight that further exploration of the HoNOS measures psychometric properties are required.

### **Outcome Rating Scale.**

The ORS was designed in the United States of America for every day clinical use by Miller and Duncan in 2000 (Bringhurst et al. 2006). The ORS is an ultra-brief, self-report outcome measure, which uses a four-item visual analogue system (Miller, 2012). The four items are, individual (well-being), interpersonal (well-being), social (well-being) and overall (well-being) (Miller & Duncan, 2000). Its simplistic design allows for easy tracking of client progress, enabling therapeutic time to be more beneficial and productive for the client. All items are measured on a ten-centimetre line and clients are asked to place a mark where they feel their well-being (for that area) best sits. Marks to the left indicate difficulties in that domain, and marks to the right indicate fewer difficulties (Miller & Duncan, 2000).

The ORS was developed using three main areas of client functioning from the OQ-45.2. These areas are individual, relational and social well-being and functioning (Miller, 2012). It is widely believed that changes in these three areas are valid and reliable indicators of successful treatment outcomes (Miller et al. 2003). Research has consistently demonstrated the validity and reliability of ultra-brief outcome measures in numerous areas of mental health and medical health care (see Radbruch et al. 1999; Zalon, 1999; Ger, Ho, Sun, Wang, & Cleeland, 1999). An important aspect of ultra-brief visual analogue outcome evaluations is their face validity (Miller et al. 2003). Their simple 'on-the-spot' administration,

scoring and interpretation are often missed with lengthy, more complex scales where true, raw client experience seems distant from the actual administration and scoring of the scale. It has been found that often in busy clinics, the clinician who assessed the client may not be the same person to complete the outcome evaluation, causing client experience to be further removed; and that is if the outcome scale is completed (Miller et al. 2003). Thus, making outcome information less reliable.

It was highlighted by Miller et al. (2006) that often outcome evaluations are not feasible in everyday clinical settings. The ORS was developed in response to this need, as an ultra-brief outcome alternative to the OQ-45.2. Miller et al. (2003) acknowledged that other more complex, lengthy, multi-dimensional assessments are valid and reliable, but often these exact factors (length of administration, cost, complexity) make them infeasible in everyday clinical settings. For example, longer more complex outcome evaluations in busy clinics can cause mayhem on schedules if people are running late. Brown, Dreis and Nace, (1999) found that clinicians believed measures that took more than five minutes to administer, score, and interpret were impractical in clinical settings. Miller et al. (2003) reviewed the ORS psychometric properties with both clinical and nonclinical populations and assessed the measures feasibility in several clinical settings. They found the measure to be valid and reliable and to be comparable to other more complex, longer measures. The feasibility of the ORS as a brief outcome evaluation is what they believe makes the ORS more 'user-friendly' and a more effective instrument.

Campbell and Hemsley (2009) studied the ORS and SRS in a psychological service for rural primary health care in Australia. They compared the measures to longer measures, such as the Depression Anxiety Stress Scale-21 and the Self-Esteem Scale. The results indicated the ORS and SRS had good reliability and concurrent validity with the more comprehensive measures. The ORS also provided evidence that the scales covered more than the four areas of wellbeing, and correlated with scales of self-efficacy, self-esteem and quality of life (Campbell & Hemsley, 2009). Bringhurst et al. (2006) also studied the ORS, but in a nonclinical student population in America. Their results found the ORS has high test-retest reliability, strong internal consistency and moderate concurrent validity. Concluding the ORS has adequate psychometric properties (Bringhurst et al. 2006). Janse et al. (2014) who studied the ORS and SRS in a Dutch mental health care setting, however, found the ORS and SRS has limited concurrent validity and adequate test-retest and internal consistency. However, they stated it was likely due to the scale needing new psychometric standards with different cultural populations. Through the research, the ORS and SRS are evidently providing support for the adoption of brief outcome measures in mental health clinical settings.

### **The Present Study**

The current study aimed to explore the performance of the ORS, an ultra-brief treatment session outcome evaluation scale, in an outpatient treatment sample within New Zealand.

Overall, the following research questions guided this study:

1. Is the ORS and SRS a reliable and valid scale within a New Zealand clinical setting?
2. Is the ORS and SRS feasible in New Zealand clinical settings? Do the ORS and SRS suit New Zealand's current expanding mental health services?
3. Does the ORS measure the same clinical characteristics as New Zealand's current Ministry of Health appointed outcome scale (HoNOS)?
4. Does the SRS benefit therapy outcomes?

The above research questions formed the following hypotheses:

1. The ORS and SRS will demonstrate adequate validity and reliability in a New Zealand clinical setting. It will have strong internal consistency but also weaker test-retest reliability and concurrent validity due to the brief nature of the scale.
2. A convergence of the ORS and HoNOSCA measures will be present, indicating construct validity. High scores on the ORS will relate to low HoNOSCA scores. It is expected that as ORS scores improve so do HoNOSCA scores. Highlighting ORS measures the same main characteristics as HoNOSCA. It will have slightly less detail due to it being ultra-brief (where HoNOSCA is complex and lengthy), however, the ORS will be similar in therapy effectiveness.
3. The SRS will correlate positively with treatment outcome on the ORS, suggesting that high treatment outcomes relates to high therapeutic alliance.



This study makes a contribution to New Zealand's mental health services and to the field of clinical psychology. No prior research was found on the ORS and SRS within a New Zealand sample. Therefore, this study provides new insight and thought into how New Zealand's health system can benefit from ultra-brief treatment outcome evaluations.

## **Chapter Two: Method**

### **Setting**

This study took place under Lakes District Health Board (DHB), whose catchment area includes Rotorua (Lakes DHB, 2017). Lakes DHB invited an exploration of the ORS, in its use within their Infant, Child and Adolescent Mental Health Service (iCAMHS). This service provides specific assessment and treatment for children and young people with mental health challenges (Lakes DHB, 2017).

This study used an archival clinical sample which was accessed through Lakes DHB outpatient database. It comprised both pre (entry) and post (exit) administrations of HoNOSCA (for Children and Adolescents) scores, and pre- and post ORS scores from iCAMHs. Included with the ORS scores were the client's Session Rating Scale (SRS) scores. The SRS measures the effectiveness of the therapeutic relationship.

### **Participants**

Lakes DHB iCAMHs provided the archival data for the clinical sample, and from them, consent was sought. Lakes DHB gave permission for the author to have access to both archival HoNOSCA data and archival ORS/SRS data that had been previously administered and recorded by Lakes DHB staff, (namely Psychologists recorded this data). Clinical clients of Lakes DHB were aware that their outcome scores could be used for research purposes, and gave approval prior to accessing their care.

The data was made up of 98 individuals including 57 females and 41 males, aged 5 to 18 years old ( $M= 12.51$ ,  $SD= 3.30$ ), with a range from

5.59 years to 18 years. According to Lakes DHB, the main ethnic categories were, New Zealand European ethnicity 69 (70.4%), 17 were Maori ethnicity (17.3%), 3 were either Other Asian and Other European ethnicities (3.1% each), 2 identified as African ethnicity (2%), and 1 identified as either Asian, Cook Island Maori or Tokelauan (1%). All of the participants were from the wider Rotorua area and accessed Lakes DHB iCAMHs.

Given that the present study is using this sample to make a comparison about the measure, a comparison to the Rotorua District's population is warranted. Rotorua Districts' population is 48% male and 52% female (Stats NZ, 2013). Of this, 37.5% of Rotorua's population is of Maori ethnicity, and 67.5% identify as New Zealand European ethnicity (Stats NZ, 2013). In comparison to New Zealand's overall demographics, 69% of people identify as New Zealand European and 14.6% identify as Maori (Stats NZ, 2013). Therefore, the present sample was not as representative of the Rotorua District's population as it was of the overall New Zealand population. The Rotorua District female and male percentages are similar to those in the present study, as is the New Zealand European sample size. However, the Maori ethnicity population is noticeably smaller in the present studies sample and represents the general New Zealand percentage of Maori ethnicity. This may reflect the finding of ethnic minority groups low access to mental health services (Rochford, 2004).

This sample cannot be considered representative of the general New Zealand community, due to the small population sample and the

young average population age of 12.51 years. However, the purpose of this study was not to generalise but to compare the ORS to the HoNOSCA in New Zealand's mental health settings. Any findings in this study regarding strengths and weaknesses of the ORS and SRS will still be useful to New Zealand's mental health sector.

### **Data Screening**

It is important to check the data set for errors prior to beginning analysis (Pallant, 2007). This ensures that there are no data errors that can affect the comparisons made between samples. Common data errors are different sample sizes and differences in the quantity of data for each participant (Lim & Ting, 2012). Screening for these errors improves the reliability of the data and is a necessary part of research as it helps prove that phenomena exist (Pallant, 2007). In the present study, data cleaning was used to reduce the number of participants and to get a quality sample for statistical comparisons. Data errors such as quantity of data for each participant was especially relevant to this study. All participants had to have completed pre- and post-administrations of both the ORS and the HoNOSCA scales to be used for analysis. Given the large number of administrations (8,908 ORS and 7,869 HoNOSCA administrations), data screening was extensive.

All data was screened for incomplete clinical data sets. Not all clients from Lakes DHB had both ORS and HoNOSCA scales used at intake and outtake. Thus, it became evident that data cleaning was necessary in order to make comparisons across measures. Individuals who did not have both measures completed at pre- and post- times were

excluded from the analysis. Only pre- and post- data was used for the analysis, if more administrations were administered they were not included in the analysis. This created a sample of 98 individuals.

Originally, a non-clinical sample of undergraduate psychology students were to be used for comparisons against the clinical data. Unfortunately, due to the quality clinical data being from iCAMHs, the non-clinical sample was deemed inappropriate for comparisons due to the average age difference. The average age of the clinical sample was 12.5 years and the non-clinical sample average age was 22 years old. This was considered inappropriate because the outcome evaluation scales administered to the two different age groups take a slightly different form, because one targeted children and the other adults. Meaningful comparisons would not have been able to be made between the groups, and differences could have been age related rather than directly related to the ORS, hence the decision to withdraw the nonclinical sample.

Once all data screening was completed, the data was loaded into IBM SPSS Statistics 24 (SPSS). SPSS descriptive statistics function was used on the data to check that there were no errors. If any errors were identified, the researcher went back into the raw data and made the appropriate changes. No errors were identified.

## **Measures**

As highlighted in Chapter One, an effective outcome scale requires good psychometric properties (Brooks, 2000). Ideally, the measures will be consistent and accurate across all administration, sensitive to change and feasible in clinical environments.

### **The Health of Our Nation Outcome Scale.**

The HoNOS (the current Ministry of Health mandated outcome evaluation assessment), is clinician rated and was developed in the United Kingdom (Te Pou, 2012). The HoNOS is used to measure the health and social functioning of people accessing mental health services in New Zealand (Pirkis, Burgess, Kirk, Dobson, Coombs & Williamson, 2005). It helps guide clinicians with treatment decision making and provides a picture of how a person's health, wellbeing, and circumstances change over time (Te Pou, 2014). In the present study, HoNOSCA was used to make comparisons to the ORS.

The HoNOSCA has 13 scales, with every scale being rated on a five-point scale; 0 meaning no problem to 4 representing severe/very severe problems (Garraida & Yates, 2000). Each scale has been constructed to be as independent as possible. The 13 scales cover a broad range of problems rather than a specific problem. These scales range from 'problems with disruptive antisocial behaviour' to 'poor school attendance'. There are four main categories grouped from the 13 scales – behaviour, impairment, symptom and social (Garraida & Yates, 2000). The HoNOS measures are normed to the New Zealand population, suggesting outcome information can be drawn directly from a New Zealand population. However, the scales limitations are ease of use (it is considered lengthy and training is required to administer it), it does not provide 'on the spot' information to help guide therapy and scoring is complicated.

As discussed in Chapter One, the HoNOS measures have mixed reviews regarding its reliability, validity and clinical feasibility (Brooks, 2000). Pirkis et al. (2005) conducted a qualitative literature summary on studies looking at the HoNOS. They found the HoNOS internal consistency, Cronbach's alpha ranges from 0.59 to 0.76, suggesting moderately high internal consistency and that the items measure similar constructs (Pirkis et al. 2005). Kisely et al. (2010) found the HoNOS to be sensitive to change in 9 out of the 12 items, and 8 out of the 13 items in the HoNOSCA version in outpatient mental health facilities in Canada. They also had an 82.2% completion rate which fell to 49% at follow-up. Brooks' (2000) study in Australia (as introduced in Chapter One) reported to find limited evidence of its psychometric properties in in-patient settings, and concluded that it was not an appropriate routine outcome measure. However, Te Pou (2012) reviewed HoNOS in in-patient and day services in New Zealand, and reported that it was easy to use, low cost, applicable to a wide range of mental health disorders and that it was psychometrically valid and reliable. The literature on the HoNOS scales suggest further research and changes are required to increase clinicians trust in the scale, its validity and reliability, and to make it more feasible in clinics.

### **Outcome Rating Scale and Session Rating Scale.**

The ORS and SRS were developed by Scott Miller and colleagues as brief alternatives to outcome evaluations (Miller & Duncan, 2000). They track client functioning and provide feedback on therapeutic relationships throughout therapy (Miller et al. 2006). It is available in 18 different languages and has numerous versions for different age groups. The ORS

and SRS are self-rated scales that are designed to provide quick, efficient information that can be used immediately (Miller et al. 2003). The ORS and SRS are cheap to administer and require minimal training. The ORS was developed as a brief alternative to the OQ-45.2 (Miller et al. 2003).

The ORS consists of four scales that assess a person's wellbeing in three general areas (overall wellbeing, individual wellbeing, interpersonal wellbeing and social wellbeing) (Miller et al. 2003). Individual wellbeing represents personal or symptomatic distress, interpersonal wellbeing measures how well the client is getting along in intimate relations, and social wellbeing indicates the level of satisfaction with social roles such as work, school, and relationships outside of the home (Miller & Duncan, 2004). Positive changes in these three areas are said to be representative of successful clinical treatment (Lambert et al. 1996). The four scales each add up to a total of 40 and have a clinical cut off score of 25. Suggesting, people who score 25 or under are said to be in enough distress to seek help (Miller & Duncan, 2004). The SRS cut off is 36. Scores that fall below 36 are considered "cause for concern" and should be discussed prior to ending the session. However, high scores (over 36) do not necessarily mean there is a strong therapeutic alliance and it is recommended to remain open to feedback (Miller et al. 2003).

The ORS is in a visual analogue format and is completed at the beginning of therapy (Miller & Duncan, 2004). Clients are asked to think back over the past week and are then instructed to place a hashtag on a 10 centimetre line on each of the four scales (individual, interpersonal, social and overall wellbeing), with low responses to the left and high



responses to the right (Miller & Duncan, 2004). The child's version uses sad faces and happy faces at either end of the 10 centre-metre line. In contrast, the SRS is completed at the end of every session, and clients are also asked to place a hashtag on four 10 centre-metre lines, that represent different aspects of the therapeutic alliance (relationships, goals and tasks, approach and method, and overall) (Miller & Duncan, 2000). For both scales, scoring is completed by determining the distance in centre-metres between the left end of the scale and the client's hashtag on each individual item. The four numbers are then added together to obtain the total score (Miller & Duncan, 2000).

In addition to hand scoring, the ORS and SRS are also available through online applications where they can be administered and measured following the same procedure highlighted above. These online applications provide real-time feedback and responses, showing if the treatment is on track. The online application measures clinically significant change by using a Reliable Change Index to identify that the change is attributable to therapy rather than chance (Miller et al. 2003). The ORS also provides session by session predictive trajectories to indicate if a client is at risk of dropping out or a negative outcome (Miller, 2012). This is achieved through comparing the ORS scores to that of similarly scoring individuals. Online applications of the ORS and SRS provide treatment outcomes, therapeutic alliance data and can show the development of the client throughout the therapeutic care (Miller, 2012). All data is coded and encrypted to ensure anonymity.

Studies have continuously validated the ORS psychometric properties (Miller et al. 2006; DeSantis, Jackson, Duncan, & Reese, 2017). The preliminary study on ORS by its creators found it to have high internal consistency as indicated by a high coefficient alpha (.93) indicating all of the items relate highly with each other. When the ORS was compared to the well-respected OQ-45.2 scale a correlation of .59 was found - a moderate correlation suggesting concurrent validity (Miller et al. 2003; Bringhurst et al. 2006; Janse et al. 2014). The pre-test and post-test t-test of the clinical and nonclinical sample scores provided strong evidence that the ORS is sensitive to change, by the clinical scores showing statistically significant changes and the nonclinical group showing no significant changes (suggesting construct validity) (Miller et al. 2003). When compared to the OQ-45.2 clinical feasibility, the ORS had an 89% administration consistency in 12 months whereas the OQ dropped from 33% to 25% in 12 months (Miller et al. 2003). The ORS has strong psychometric properties overall, but has less reliability than comprehensive scales due to its brevity (Bringhurst et al. 2006; Janse et al. 2014).

## **Research Designs**

This study used a quantitative, evaluation approach to investigate the ORS and SRS reliability, validity, and feasibility in New Zealand clinical settings. It looked at ORS's psychometric properties and considered the overall feasibility of ORS in the current New Zealand clinical context through comparing it to clinical HoNOSCA scores. Quantitative data was

collected in the form of clinical outcome evaluation scores (HoNOSCA scores and ORS scores) administered by Lakes DHB.

A nonexperimental correlational design approach was taken in collecting quantitative data. It employed a nonrandomised and repeated assignment of participants, where participants information was gathered following their treatment at both pre- and post- sessions.

Descriptive methods were used for quantitative data. Descriptive methods are often used to test hypotheses and to look at occurrences of behaviour (Mitchell & Jolley, 2013). It is also frequently used to examine relationships between variables to see whether they correlate. The present study looked at the relationships between ORS/SRS and HoNOSCA scores as well as the scores and their demographic information.

Mitchell and Jolley (2013) state that the main advantage of descriptive methods is being able to use them when variables cannot be manipulated for both ethical or practical reasons. In the current study, variables were not manipulated for both ethical and practical reasons. Ethically, the people who had accessed Rotorua mental health services had not consented to any outside contact with the author. Indeed, the author is not qualified to work in that capacity with the clients, so serious harm in the form of psychological distress to clients could have been caused by the intrusion. Importantly, this study is dealing with young, vulnerable children whose wellbeing is impacted, hence the manipulation of variables could increase the risk of harm to themselves and others around them. Practically, as illustrated throughout this research, clinicians are busy people, intruding on their work is not practical or feasible. Given

their many daily tasks, they may simply not have time to manipulate variables related to ORS/SRS or HoNOSCA evaluations.

## Procedures

Lakes DHB Mental Health and Addiction Services staff, namely, psychologists gathered the ORS, SRS and HoNOSCA clinical sample data. All staff had previously been trained in the administration of both ORS and HoNOSCA. The ORS was introduced to Lakes DHB to provide more effective and beneficial therapy. Not only freeing up valuable time, resources and staff, but also giving off similar levels of information compared to more comprehensive scales. During this time, Lakes DHB administered both the ORS and HoNOSCA. All ORS data was recorded into the online “My Outcomes” database programme, and all HoNOSCA data was entered into the Lakes DHB database. However, for the purpose of this study, only pre- and post- data was used for comparisons, and all other data was used for ORS psychometric properties. If any sample sets did not have pre- and post- scores they were excluded from the analysis. Lakes DHB gained consent from all service users for their ORS and HoNOSCA scores to be used for research and statistical purposes.

Table 3

### *Demographics of the Sample*

Items	<i>N</i>	<i>M</i>	<i>SD</i>
Sample	98	12.51	3.30
Males	41	12.54	3.51
Females	56	12.64	3.15

## **Data Analysis**

A statistical analysis program called SPSS Statistics for Windows, Version 21.0 was used to analyse all data. SPSS is frequently used by Psychologists and other social scientists due to its user-friendly statistic packages (Coakes & Ong, 2011).

A pretest-posttest treatment, nonexperimental, correlational design approach was taken in collecting quantitative data. This is when variables exist naturally and are not directly manipulated (Pallant, 2007).

Correlational techniques describe the strength and direction of the relationship between two variables (Mukaka, 2012). In the present study, different levels of wellbeing exist and they occur whether we are present and measuring it or not. This research aimed to explore the ORS and SRS, to measure if the ORS accurately measures change in wellbeing in clinical therapeutic environments. For this reason, correlational techniques were appropriate and were used to measure changes in the ORS and SRS to make comparisons to HoNOSCA. It is important to note, that although these correlations can show that a relationship exists between two variables, such as the ORS scale measuring similar constructs to the HoNOS scale, it does not imply causality.

## **Ethical Considerations**

It is of the utmost importance for all research to take extensive considerations regarding ethical principles. Ethical considerations are paramount in every study. This is especially relevant when the study uses human participants to ensure the privacy, safety, health and welfare of all participants. Research on sensitive topics can present further challenges,

due to outcome evaluations containing personal information, the present study can be considered a sensitive topic due to its access to sensitive information. This research was reviewed and approved by the School of Psychology Ethics committee.

Munford, Sanders, Veitch, and Conder (2008) highlight the importance of researchers remaining impartial and not forcing their morals and beliefs on others. This helps create a safe research environment and ensures the research represents the participant's views fairly, and accurately. Specifically, regarding outcome evaluations in clinical settings, it is important that clients feel secure to express their wellbeing and thoughts openly, and honestly (Code of Ethics, 2002). This creates an environment where reliable and accurate outcome evaluation data can be recorded. Ethically, it is a clinician's responsibility to develop this open therapeutic relationship. In the present study, the researcher had no direct control over how data was obtained. Instead, it was taken on good faith that clinicians followed ethical procedures as per the Psychology Code of Ethics when administering outcome evaluations.

### **Ethics of Care.**

In the current Psychology Code of Ethics (2002), it highlights the basic ethical expectation of ethics of care. Researchers and clinicians alike have an obligation to others. This represents the obligation to others in connectedness in relationships, cooperation, and communication. A part of this obligation to others is to identify why the research is important and who is going to benefit from it (Code of Ethics, 2002). In the present study,

Lakes DHB will benefit from seeing how the ORS has fitted into their services. It will also benefit further health boards who are looking for a brief outcome evaluation alternative that is also valid and reliable in a New Zealand clinical setting. Ultimately, this research may open up a discussion on outcome evaluations in New Zealand's mental health services. Arguably, this could benefit everyone.

Munford et al. (2008) identify three key areas of ethics which protect the interests of participants when conducting research with human participants. These are, non-maleficence, beneficence, and justice. These three areas are the cornerstone of ethically sound research where harm to participants is avoided, informed consent is gained, and research data is used appropriately.

*A risk of harm* was not a major ethical concern in the present study. Clinical participants were in a safe, secure, therapeutic setting when outcome evaluations were administered. All of their recorded information is kept in a secure system held through Lakes DHB. When Lakes DHB gave the researcher access to the data, all identifying information (names, NHI numbers) had been removed. The researcher was the only external person outside of Lakes DHB to have access to the data and the data was kept on a password secured computer.

*Informed consent* is an important ethical consideration with research (Code of Ethics, 2002). Informed consent needs to be voluntarily obtained, it cannot involve deception, and the researcher needs to ensure the participant is fully aware of the research implications and their involvements. The Code of Ethics for Psychologists working in

Aotearoa/New Zealand speaks directly to this in the value statement on informed consent: “Psychologists recognise that obtaining informed consent from those with whom they are working is a fundamental expression of respect for the dignity of persons and peoples” (Code of Ethics, 2002, p. 8). In the present study, informed consent for the clinical sample was obtained through Lakes DHB prior to their taking part in therapy.

Under the Official Health Act 1956, the Health (Retention of Health Information) Regulations 1966, as well as the Official Information Act 1982, in New Zealand, all data and information obtained in New Zealand’s health services become the property of the government and can be used for statistical and research purposes; in accordance with the Health Information Privacy code 1994 (Ministry of Health, 2012). All information regarding the clinical sample was randomly enumerated and all information was kept confidential and secure. An agreement between the researcher and the key stakeholders at Lakes DHB was formed where all data that was given was only to be used for this research purpose and all data would be erased after ten years.

It is ethically important to only use the research data for the agreed purpose (Code of Ethics, 2002). Misuse of research data and information is an ethics issue that needs to be closely monitored. The Code of Ethics (2002) which psychologists follow, highlights the importance of informed consent and ensuring all information and data that is used was obtained under informed consent. Upholding to this agreement ensures the safety and wellbeing of all participants. Information for the current study was



collected to gain a better understanding of the ORS and SRS, regarding the measures validity, reliability and feasibility in New Zealand clinical settings. All information gathered was only used for this purpose.

### **Relationships.**

The information gathered for the current study did not directly deal with people. However, confidentiality was still a crucial ethical consideration. Everyone has a right to privacy, and it is a researcher's obligation to protect this right (Munford et al. 2008). The clinical sample understood their information could be used for research and statistical purposes and were explained the limits to their confidentiality.

To protect the clinical samples confidentiality several measures were taken. All data was stored securely and only the researcher had access to it. All data was anonymized, where the clinical sample was randomly coded by a Lakes DHB data analyst, to ensure all NHI, names or identifying information was taken away. This thesis holds no identifying information to protect the privacy of all participants.

The right to withdraw at any time is an ethical consideration that needs to be clearly highlighted to all participants (Code of Ethics, 2002). For this research, it is important to note that clinical participants could not withdraw due to their data being historical and consent being given through their admission.

### **Cultural Considerations.**

Culture is of the utmost importance in all research, because it can influence outcomes and have major impacts on communities (Munford et

al. 2008). It is especially important in New Zealand given New Zealand's indigenous Maori heritage. Maori as a group experience relative socioeconomic disadvantage compared with other communities across a range of life domains (education, health, crime and justice) (Chapple, 2000). Western health models are often used in New Zealand and these do not always match traditional Maori worldviews of health. Te Whare Tapa Wha is a Maori model of health developed by Durie in 1982 (Rochford, 2004). It acknowledges traditional Maori understandings of health and is still valid today. In the current study, culture may influence what an individual sees as good or bad wellbeing and the outcomes may have been impacted by the health models accessed. Culture is especially relevant to the current study as a large portion of the Rotorua population are Maori (Stats NZ, 2013). Rotorua's 2013 Census data found 37.5% identify as Maori compared to 14.9% for all of New Zealand (Stats NZ, 2013).

New Zealand's Ministry of Health works alongside the Maori Health Group to ensure Maori health values are recognised in our health systems, and also to ensure the principles of the Te Tiriti o Waitangi (the Treaty of Waitangi) are upheld (Ministry of Health, 2017). If Maori health values are actively acknowledged and used in practice, it is believed that all New Zealander's will have equal and beneficial health care (Rochford, 2004). New Zealand's current health care system is dominantly based on Western European values rather than holistic Maori health values. It is reported that often Maori are not well represented in New Zealand's mental health services, causing problems in Maori accessing care and

feeling heard (Rochford, 2004). For this reason, it is important that all healthcare practitioners are culturally competent and willing to further their cultural awareness.

Cultural competence, can be understood as a constant 'work in progress', which is defined as, "having the awareness, knowledge, and skill, necessary to perform a myriad of psychological tasks that recognises the diverse worldviews and practices of oneself and of clients from different ethnic/cultural backgrounds" (New Zealand Psychologists Board, 2011, p. 4). Apart of being culturally competent is being safe. Protecting people's cultures and ensuring you understand the cultural implications research can have.

It was important for this research to acknowledge the potential impacts on Maori. The present research explored the ORS/SRS and the measures validity, reliability and feasibility in New Zealand clinical settings, and this included how it represents and relates to Maori service users. Currently, there is no research on the ORS or SRS in New Zealand, let alone on the Maori population. The outcome of this research will be able to be used to guide future use of the ORS and SRS with a New Zealand population. Including if it measures a change in mental health services for Maori users.

## **Chapter Three: Results**

This chapter describes the results of the present study regarding each of the three hypotheses outlined in the introduction. It begins with the descriptive statistics for each of the measures (ORS, SRS, and HoNOSCA). This is followed by systematically working through each of the hypotheses. Finally, the overall feasibility of the ORS and SRS will be discussed.

### **Descriptive Statistics**

Table 4 displays the means and standard deviations for the total ORS, HoNOSCA and SRS samples. As expected, the ORS's pre- and post- scores indicate positive change following therapy. The present studies ORS clinical intake and outcome scores were slightly higher than those reported in the preliminary ORS reliability and validity study (Miller et al. 2003). This suggests that New Zealand's population presents with slightly higher wellbeing and possibly need slightly different norms than those created by the American population. It could also reflect the samples ages, acuity and understanding of the scales. Total pre- and post- HoNOSCA means also reflect positive change following therapy. The mean and standard deviations for pre- and post- HoNOSCA scores are comparable to those reported in a British review of HoNOSCA (Garraida & Yates, 2000). The SRS means increased, indicating that most therapists provided a positive working alliance from the beginning of treatment, and as treatment progressed that alliance strengthened.

Table 4

*Comparison of ORS, HoNOSCA and SRS Pre- and Post- Data*

Instrument	<i>n</i>	<i>M</i>	<i>SD</i>
<i>ORS</i>			
<i>Pre-</i>	98	23.267	7.7065
<i>Post-</i>	98	30.295	8.7538
<i>HoNOSCA</i>			
<i>Pre-</i>	97	12.08	4.960
<i>post-</i>	97	7.26	4.931
<i>SRS</i>			
<i>Pre-</i>	95	32.42	6.08
<i>post-</i>	93	35.37	5.43

Miller et al's. (2003) normative data suggest that individual scores do not differ due to age or sex. In the present study, no significant differences in HoNOSCA and ORS intake scores were found between sex. However, there was a significant correlation between ORS scores and age. This is further broken out below. Table 5 displays the means and standard deviations for the ORS scores by gender.

Table 5

*Gender Comparisons of ORS Means and Standard Deviations*

Gender	Sample Size	<i>M</i>	<i>SD</i>
Males	41	23.405	7.6782
Females	56	23.095	7.8439

In contrast to Miller et al's. (2003) study, a moderate negative correlation between age and total pre- ORS scores was found, ( $r = -.463$ ,  $p, .01$ ). This suggests that as participants get older, their ORS scores get lower, or as ORS scores increase age decreases. This likely reflects participants having better insight and awareness of their wellbeing levels as they get older.

HoNOSCA scores were at large not impacted by age. However, HoNOSCA measure 13 'poor school attendance' had a weak positive correlation with age ( $r = .296$ ,  $p, .05$ ). HoNOSCA was not expected to correlate with age given that it is clinician rated. However, measure 13 relates to school aged children, given that the average participant age is 12.5 years, this weak positive correlation is not surprising. Likely suggesting, that the average participant in the sample has poor school attendance due to their challenges with wellbeing.

### **Ethnicity and ORS.**

Maori and New Zealand European ORS scores are comparable. Maori have slightly lower intake scores and roughly the same overall outcome scores compared to New Zealand European ORS scores. Table 6 displays the means and standard deviations for both ethnic groups. Both ethnic groups pre-and post- scores had significant moderate to strong positive correlations with their pre- and post- scores, New Zealand European pre- and post- correlation was ( $r = .568$ ,  $p, .01$ ), and Maori pre- and post- ORS correlation was ( $r = .676$ ,  $p, .01$ ).

Table 6  
*Maori and European Comparisons*

Measures	<i>N</i>	<i>M</i>	<i>SD</i>
Maori ORS			
<i>Pre-</i>	17	21.765	8.3904
<i>Post-</i>	17	30.176	6.6950
NZ European ORS			
<i>Pre-</i>	69	23.335	7.5481
<i>Post-</i>	69	30.687	9.0812

### **Correlations**

Correlations were used in this study to measure and compare relationships between the ORS variables, SRS variables, and the HoNOSCA variables. Correlation describes the strength and direction of a relationship between two variables (Pallant, 2007). Correlations range from +1 which is a perfect positive linear relationship, to -1 which is a perfect negative linear relationship. Correlations were taken between, pre- and post- ratings of each individual item and overall on the ORS and HoNOSCA scales, as well as with each outcome scales age, gender and ethnicity. A correlation between ORS scores and SRS scores was also computed.

### **Hypothesis One**

Hypothesis one aimed to demonstrate the reliability and validity of the ORS and SRS. It was expected that the ORS will have strong internal consistency but weaker reliability and concurrent validity. Data to support

Hypothesis One was limited due to the sample data being patchy. However, there was evidence of the ORS reliability and validity. Each individual ORS factors correlated moderately to strongly with each other, as shown in Table 7. This indicates that the ORS has some internal consistency reliability, and suggests that the ORS could be viewed in terms of a global measure of distress instead of a scale looking at specific areas of wellbeing. Additionally, pre- and post- ORS scores were correlated to test the test-retest reliability of the measure. Pre- and post-total ORS scores had a moderate to strong positive correlation ( $r = .543$ ,  $p < .01$ ). This could reflect the measure being sensitive to small changes, reinforcing the global measure of distress idea. In terms of validity, there was a weak negative correlation ( $r = -.284$ ,  $p < .01$ ) between means of the measures, suggesting some evidence of concurrent validity. A negative correlation was expected given that as HoNOSCA scores improve the score decreases, whereas when the ORS scores improve the score increases. Lastly, the positive increase in ORS scores, indicates that it is sensitive to change.

Table 7

*Correlations Matrix Between Individual ORS Items*

Items	ORS 1	ORS 2	ORS 3	ORS 4
ORS 1				
<i>Individually</i>	1			
ORS 2				
<i>Interpersonally</i>	.793**	1		
ORS 3				
<i>Socially</i>	.743**	.617**	1	



<i>ORS 4</i>				
<i>Overall</i>	.933**	.764**	.754**	1

\*\* . Correlation is significant at the 0.01 level (2-tailed).

## Hypothesis Two

Hypothesis two aimed to measure the convergence of the ORS and HoNOSCA measures. It was expected that as ORS scores improve, so do the HoNOSCA scores. Pre- and post- HoNOSCA scores had a weak, positive linear relationship ( $r = .370, p < .01$ ), individual items on the pre- and post- HoNOSCA scale mainly correlated strongly with their own sample set. However, when individual post- scores were compared to individual pre- scores very weak, nonsignificant positive correlations were revealed. This suggests that individual HoNOSCA items are measuring slightly different constructs. Table 8 displays the correlation matrix for pre- and post- ORS and HoNOSCA totals.

Table 8

*Correlation Matrix for Pre- and Post- ORS and HoNOSCA Totals*

Items	<i>Total ORS Pre</i>	<i>Total ORS Post</i>	<i>Total HoNOSCA Pre</i>	<i>Total HoNOSCA Post</i>
<i>Total ORS</i>				
<i>Pre-</i>	1			
<i>Post-</i>	.543**	1		
<i>Total HoNOSCA</i>				
<i>Pre-</i>	-.214*	-.216*	1	
<i>Post-</i>	-.105	-.284**	.370**	1

\*\* . Correlation is significant at the 0.01 level (2-tailed).

\*. Correlation is significant at the 0.05 level (2-tailed).

The significant relationship between overall post ORS and HoNOSCA scores initiated a further breakdown of the specific item relationships between each scale. The following results reveal the correlational relationship between items on each measure.

The correlations indicated that the ORS measure is comparable to the HoNOSCA measure in three areas (Overall wellbeing, individual wellbeing and interpersonal wellbeing). Not all correlations were found to be significant, and not all items of each measure correlated. In fact, only four out of 13 HoNOSCA items correlated, and the ORS item 3 'Social Wellbeing' did not significantly correlate with any of the 13 pre-and post HoNOSCA measures.

HoNOSCA measure 2 represents problems with over-activity, attention or concentration. HoNOSCA measure 2 had a weak negative correlation with pre-ORS 'Overall Wellbeing' ( $r = -.229, < p, .05$ ), post ORS 'Individual Wellbeing' ( $r = -.264, < p, .01$ ), and post ORS 'Interpersonal Wellbeing' ( $r = -.255, < p, .05$ ). HoNOSCA measure 3, measures non-accidental self-injury and had a weak negative correlation with post ORS 'Individual Wellbeing' ( $r = -.242, < p, .05$ ), and post ORS 'Overall Wellbeing' ( $r = -.286, < p, .01$ ). HoNOSCA measure 12 assesses 'problems with family life and relationships', and had significant weak negative correlations with pre-ORS 'Interpersonal Wellbeing' ( $r = -.231, < p, .05$ ), pre- ORS 'Overall Wellbeing' ( $r = -.200, < p, .05$ ), post ORS 'Individual Wellbeing' ( $r = -.208, < p, .05$ ), post ORS 'Interpersonal Wellbeing' ( $r = -.264, < p, .01$ ), and post ORS "Overall Wellbeing' ( $r = -.205, < p, .05$ ). The last HoNOSCA measure that significantly correlated with the ORS is Post

HoNOSCA measure 13, which measures 'poor school attendance'. Post measure 13 had a weak negative correlation with pre-ORS Individual Wellbeing ( $r = -.283, < p, .01$ ), and pre- ORS Overall Wellbeing ( $r = -.276, < p, .01$ ).

Overall, these results indicate that there is some convergence of the ORS and the HoNOSCA measures outcome factors, with a particular convergence with the ORS 'Overall Wellbeing' item. This provides evidence of the number of constructs that the ORS item "Overall Wellbeing" measures, highlighting its global measure capacity. It also highlights that both scales are reflecting a therapeutic change, and measuring the constructs that they say they are.

### **Hypothesis Three**

Hypothesis three aimed to demonstrate that high session ratings are indicative of positive treatment outcomes. It was expected that SRS scores will positively correlate with ORS scores. The SRS is used to measure the working alliance between the client and the therapist in clinical settings (Miller & Duncan, 2000). Total SRS pre- and post- scores indicated that the SRS improved over the course of treatment. Total pre- and post SRS scores had a moderate positive correlation ( $r = .401, < p, .01$ .) Both pre- and post- SRS scores significantly correlated with pre- and post- scores of the ORS at a .01 two-tailed significance level. These positive correlations suggest that as the SRS scores increase so do the ORS scores, particularly the overall wellbeing measure ( $r = .349, < p, .01$ ) and that the SRS and ORS have good internal consistency. This reinforces that positive therapeutic relationships are beneficial to therapy

outcomes. Table 9 displays the SRS means and standard deviations, and Table 10 displays the correlations between the SRS scores and ORS scores. There were no significant correlations between age and SRS scores.

Table 9

*Session Rating Scales Means and Standard Deviations*

Items	<i>M</i>	<i>SD</i>	<i>N</i>
<i>Total SRS</i>			
<i>Pre</i>	32.42	6.08	95
<i>Post</i>	35.37	5.43	93
<i>Total ORS</i>			
<i>Pre</i>	23.26	7.70	98
<i>Post</i>	30.29	8.75	98

Table 10

*Correlation Matrix of SRS scores and pre- and post- ORS scores*

Items	<i>Total SRS</i>	<i>Total SRS</i>	<i>Total ORS</i>	<i>Total ORS</i>
	<i>Pre</i>	<i>Post</i>	<i>Pre</i>	<i>Post</i>
<i>Total SRS</i>				
<i>Pre-</i>	1			
<i>Post-</i>	.401**	1		
<i>Total ORS</i>				
<i>Pre-</i>	.385**	.267**	1	
<i>Post-</i>	.293**	.535**	.543**	1

\*\* . Correlation is significant at the 0.01 level (2-tailed).

### **Feasibility of the ORS**

Although directly measuring feasibility was outside the scope of this study, the clinical data obtained shows an extent of the compliance rate. In the present study, HoNOSCA discharge or 3 monthly reviews, were often incomplete or not administered. This could reflect some of the challenges therapists face with administering HoNOS measures, such as length of administration. The present study reveals the ORS data to also not be recorded perfectly, however, even though HoNOSCA has more administrations (2,783 compared to ORS 1,182 from 2014 to 2016) overall it has slightly more incomplete administrations, and even fewer administrations than the overall ORS in the raw data. This indicates that the feasibility of the HoNOS family of measures is inconsistent and suggests the scale has challenges regarding feasibility.

## **Chapter Four: Discussion**

The findings of this study provide evidence to support the adoption of the ORS and SRS as ultra-brief treatment outcome measures in clinical settings. This is suggested by the ORS effectively measuring change in a New Zealand outpatient clinic, being comparable to HoNOSCA (on certain measures), and through the SRS being a beneficial tool to facilitate discussion about therapy.

The ORS and SRS have obvious face validity for clients and evidence of concurrent validity with HoNOSCA. This is suggested through the ORS improving overall from pre- scores to post scores, by the ORS being significantly correlated with HoNOSCA measures on several appropriate scales, and by the ORS getting similar mean scores to that of the original ORS study by Miller et al. (2003). The positive correlation between the ORS scores and SRS scores, reinforces that therapeutic alliance is an important part of the process of change in therapy.

Indicating, the SRS can help determine the appropriateness of therapy quickly, and perhaps, create more effective therapy outcomes. The gaps in the present study's HoNOSCA data may show there are challenges with feasibility, which suggest a need for a briefer outcome evaluation alternative for routine use, when clinics are running short of time, but still require useful outcome information, without cutting corners.

This chapter explores the meaning of the findings. It compares the findings to the hypotheses and to current literature. Practical implications and study limitations of this research will be discussed. The chapter

concludes with a discussion on possible future directions for research and some concluding thoughts.

## **Hypothesis Testing**

### **Hypothesis One.**

The findings of this study begin to support hypothesis one and indicate that the ORS and SRS have adequate psychometric properties, and both the ORS and SRS are sensitive to change. Notably, the results provided evidence that the ORS and SRS were reliable and valid measures in a New Zealand population. However, further analysis is required to measure the full extent of the ORS and SRS validity and reliability due to challenges with the present studies data. Mean scores for the ORS were similar to that of the original study by Miller et al. (2003) and significantly changed overall as therapy progressed. This suggests that the ORS was picking up on change happening in therapy sessions.

Another strong factor in the ORS reliability is the moderate to strong positive correlations between each of the four ORS items found in this study. Following suit, the SRS also has significant weak to moderate internal consistency. This finding has consistently been found in previous research (see; Miller et al. 2003; Bringhurst et al. 2006). It is frequently found that smaller measures have lower internal consistency. However, these moderate to strong positive correlations indicate the ORS and SRS high degree of internal consistency. Bringhurst et al. (2006) suggest the high internal consistency of the ORS could mean the ORS could be thought of as a global measure of distress, rather than using it for its subscale dimensions. This finding was recreated in the present study.

There was some evidence of concurrent validity between the ORS and HoNOSCA, however there are very few significant correlations between the two measures. Given the shorter nature of the ORS compared to the HoNOSCA, and given the scales differ slightly this was expected. The ORS provided some evidence of construct validity through it being sensitive to change in pre- and post- ORS and SRS scores.

### **Hypothesis Two.**

Hypothesis two was supported to an extent, as not all items on each measure converged. Overall post scores on the ORS had a significant weak negative correlation with overall post HoNOSCA scores. A negative correlation was expected given that as HoNOSCA scores improve the score decreases, whereas when ORS scores improve the score increases. However, when the measures were further broken down, specific items between the measures did not show the same level of significance. In fact, the ORS item “Social Wellbeing” did not significantly correlate with any HoNOSCA items. This may mean that what clients rated as their social wellbeing, was not covered in the HoNOSCA measure. However, the items that did significantly correlate, all correlated with the ORS item 4 “Overall Wellbeing”. This was expected given that overall wellbeing encompasses all aspects of wellbeing and is likely to overlap with many HoNOSCA measured items.

Further broken down, HoNOSCA measure 2 (over-activity, attention or concentration) significantly correlated with ORS’s Overall Wellbeing, Individual Wellbeing, and Interpersonal Wellbeing. Upon looking at this closer, these items would be expected to correlate due to over-activity



being an individual dominated behaviour that has repercussions for other people they interact with, hence the relationship with interpersonal wellbeing. HoNOSCA measure 3 (non-accidental self-injury) had significant relationships with individual wellbeing and overall wellbeing. Once again, this relationship was anticipated due to non-accidental self-injury being a heavily individual driven behaviour. The relationship with overall wellbeing further illustrates the extent of what the item measures. HoNOSCA measure 12 (problems with family life and relationships) revealed significant relationships with ORS items interpersonal wellbeing, individual wellbeing and once again, overall wellbeing. Problems with family life and relationships occurs between people and is therefore an interpersonal wellbeing challenge, and with it, individual challenges can occur. The final HoNOSCA measure that significantly correlated with the ORS was HoNOSCA measure 13 (poor school attendance). This measure correlated significantly with the ORS individual wellbeing and ORS overall wellbeing. This is not a surprising finding given that the average age of the clinical sample is 12.51 years. This means the average sample age is a school aged child who is seeking mental health support. It is not farfetched that a child's school attendance would be affected by their mental health challenges.

These findings suggest that both the ORS and HoNOSCA are measuring some similar constructs, however it is only at a weak significance level. It would be too much of a claim to state that the ORS measures exactly the same constructs that the HoNOSCA scale measures. Especially given the broad subjective nature of the ORS. Given

the difference in purpose of both measures and the brevity of the ORS, it is not unusual the findings suggest they are not measuring exactly the same constructs.

Both measures were found to measure change in clients throughout therapy, however the ORS tracks this change more frequently (every session rather than pre- and post and 3 monthly intervals while in the service). This finding supports the final part of hypothesis two that the ORS will measure significant clinical change as HoNOSCA measures.

### **Hypothesis Three.**

Hypothesis three was supported. There was a significant moderate positive correlation between the ORS and SRS. Highlighting that facilitating open discussions about sessions, provides a solid foundation for positive treatment outcomes. It may also suggest, that clinical sessions with strong therapeutic alliance can help predict therapy outcomes. This finding was also found in the preliminary study on an American population (Miller et al. 2006). Miller et al. (2006) explain the positive connection between the ORS and SRS scores, by stating that it is not solely a score of therapeutic alliance, rather it is a method of opening up discussions around what works for the client and what does not work. They suggest it is these conversations where the “magic” of therapeutic change happens. The SRS enables clients to use their own perspectives on therapeutic alliance, rather than a therapist relying on their own. Shaw and Murray (2014) state that a client’s perspective on therapeutic alliance is a better indicator of outcome than the therapist’s perspective. The present study’s findings allude to this.

### **Feasibility.**

Although feasibility was somewhat out of the scope of this research, a major part of proving the ORS's feasibility in clinical settings is the likelihood of the measure being used (Miller, 2012). Miller et al. (2006) put an emphasis on a tool's feasibility and state that if a tool is not feasible in a clinic, its strong psychometric properties are useless. Notably, the ORS data was solidly recorded by Lakes DHB, and often more so than the current New Zealand mandated outcome evaluation. This may suggest either that the ORS provided clinicians with the information they required, saw most beneficial, or that clinicians found the tool more accessible in their clinic setting over the HoNOS tools.

An important area of feasibility in New Zealand, is the cultural responsiveness of the measure. Proving that the ORS responds to New Zealand's culture effectively is beyond looking at data sets and is likely something that needs some qualitative input. It is worthwhile to mention that the HoNOS measure developers state the HoNOS family of measures does not cater to spiritual or existential difficulties (Te Pou, 2012). Given New Zealand's indigenous culture has numerous spiritual beliefs, you would expect a government mandated scale to cover this important element of wellbeing. The ORS is far too brief to be able to cover the spiritual wellbeing element of New Zealand's culture. Although, the broad client-rated wellbeing scales do allow room for interpretation of what wellbeing means for each individual. The results of the present study found people who identified as New Zealand European and New Zealand Maori had similar mean scores and both ethnicities' individual scores

provided evidence of meaningful change. Maori intake scores were slightly lower than those of New Zealand European but both ethnic groups outcome scores were relatively the same. This finding indicates that the Maori population in the clinical sample improved over therapy, and that perhaps the measure responded to their wellbeing needs.

Recent survey data highlights that there is a rise in Maori mental health illness, mixed with low rates of Maori accessing services (New Zealand Mental Health Survey, 2014). This finding suggests a need to incorporate Maori health values into our services and utilising tools that allow all voices to be heard. It is important to mention that some researchers such as Bennett, Flett and Babbage (2016) suggest that effectively incorporating Maori into therapy does not just mean adapting Western models, but creating and utilising a model developed specifically for Maori. As mentioned in Chapter Two, incorporating Maori models of health, such as Te Whare Tapa Wha may be a way forward in authentically including Maori health outcome aspects into mental health care (Rochford, 2004). Changes such as these, may develop an environment where the common factors and process of change can occur freely.

## **The Literature and the Present Study**

### **Psychotherapy.**

The current study further strengthened Janse et al's. (2014) statement of psychotherapy being effective. Overall, the results indicated that clients were better off after therapy suggesting that meaningful change occurred, and ultimately, that the ORS measures change.

As discussed in Chapter One, having an understanding of what psychotherapy encompasses, is important in understanding therapeutic outcomes. By being able to identify what the current literature says, what is happening in therapy, and within therapy services, outcome evaluations can be used more effectively to measure known areas of meaningful change. An important factor highlighted in the current study is that mental health services are changing in numerous areas, such as how the service is delivered (there is a massive trend towards integrative therapies), and the amount of people accessing services, which is causing stricter time limits (New Zealand Mental Health Survey, 2014; Miller et al. 2003; Wampold & Imel, 2015). When these factors are compiled together it becomes increasingly apparent that there needs to be options of outcome evaluations to best suit not only the clinician, but also the service user.

Having options of outcome evaluations gives clinicians choices and increases the likelihood of outcome evaluations being administered. According to Janse et al. (2014) this will also heighten the chances of successful therapy outcomes. Janse et al's. (2014) research revealed that positive therapy outcomes were heavily influenced by measuring outcome and remaining "on the same page" as the client. Miller et al. (2003; 2006; 2013) stated that the SRS was one tool that helps facilitate open discussion about the therapy process for clients. The present studies findings revealed high SRS scores with high ORS scores, possibly supporting Janse et al's. (2014) and Miller et al's. (2003) research of positive outcomes coming from open discussions.

Something that became overwhelmingly apparent across the literature was the heavy emphasis placed on an operational era of outcome evaluations, and then suiting the needs of the clinician and/or stakeholder/policy makers (Miller et al. 2006). Articles frequently discussed the time challenges clinicians have and their opinions of the scales. This fundamental thinking appears problematic and needs to be flipped to where the client's needs are at the centre and everything else stems from them. Thoughts such as, "what outcome evaluation will best suit my client's needs", and "where does this sit with the client?", need to be at the front of all outcome evaluation administrations. An important thought for clinicians to move forward with is asking what the evaluation means for the client. If a client believes the outcome evaluation has a purpose, perhaps therapy will be more meaningful as suggested by the research on common factors of therapy, specifically a positive expectancy trait (Cohen et al. 2015). This change would reflect the trend towards finding a specific treatment that is best suited to a client's set of challenges, and perhaps would engage some common factors in therapy. If the client's interests are at the heart of outcome decisions, useful information and results may start to flow.

As discussed in Chapter One, psychometric properties are a necessary component of outcome measures, but they also need to be highly feasible (Miller & Duncan, 2000; Te Pou, 2012). The results of the present study indicated that the ORS and SRS are reliable ultra-brief treatment measures in therapy, that stand out due to the high feasibility of the measures. Miller and Duncan (2003) state that the ORS has valid and

reliable psychometric properties, but the statistical strength of them can never be as strong as a comprehensive outcome evaluation. Suggesting that an ultra-brief outcome scale such as the ORS could be used routinely during sessions, and other more comprehensive scales can be used at pre- and post- administrations. This would allow comprehensive data to be recorded without impacting the day to day functioning of the clinic. By using the ORS in conjunction with other outcome evaluation scales, the inter-rater reliability would improve due to it no longer only being a sole administered scale.

### **Outcome Measures**

Common factors such as expectation for improvement, persuasion, warmth and attention, understanding and encouragement are well known factors that can impact therapy outcomes (Lambert, 2004; Wampold & Imel, 2015). The current study highlighted the strength of open discussion in therapy and the importance of a therapeutic alliance as a measure for therapy outcomes. Results indicated that when clients rated highly on the SRS they also rated higher on the ORS. This supports that when a client rates the therapeutic relationship highly, therapy is more likely to create meaningful change. Other components of the SRS look at the goals and topics discussed, whether they were relevant, how the client found the approach and method of the therapist, and if the client feels the therapist is a 'good fit' for them. All of these factors in the SRS relate closely to Lambert's (2004) identified common factors. This suggests that outcome evaluations should not only strictly assess clients' functioning, but also open up dialog on what is working in therapy as well as what is not

working. Relationships such as these, will help build therapeutic alliance, and enable clinicians to produce meaningful change and monitor risks such as dropouts.

Duncan and Miller (2000) highlighted that early signs of change are important and incorporating outcome information in treatment are linked to overall positive therapy outcomes. The overall findings of the current study showed positive treatment outcomes. However, it is unknown if this was due to the routine ORS use. New Zealand's current outcome measures (HoNOS family of measures) are usually assessed at the initial consultation as well as at discharge and at random points such as 3 monthly intervals. The promising results connecting the SRS and positive outcomes, suggest outcome evaluations should be used more frequently to see if meaningful change is occurring, and if it is not, changes can be made.

Progress, such as expanding outcome evaluations to incorporate SRS can be viewed as a part of the ever-expanding field of routine outcome measurement. As discussed in Chapter One, the purpose of outcome evaluations has developed over time. From initially being used to measure effectiveness with no standardisations, to being used routinely in clinics with clear standards and outcome goals (Wampold & Imel, 2015). The movement towards ultra-brief outcome evaluation alternatives, highlight the current clinical time where clinicians are engaging in more brief therapy, and clients are treated in a standardised manner specifically for diagnoses, rather than a broad array of challenges. The results of the current study suggest that ultra-brief outcome evaluations are a possible



way forward in measuring meaningful change and ensuring that therapy provides positive outcomes in the most accessible format possible.

### **Practical Implications**

The ORS and SRS have clear face validity and it is this simplicity that makes ultra-brief outcome evaluations more feasible in clinical environments. Literature frequently identifies that lengthy, complex outcome measures are likely to be dropped when clinicians are busy due to the tools taking up valuable clinical time, and not necessarily providing 'on the spot' guidance for therapy (Miller et al. 2006). The patchy HoNOS data from the current study and recent figures released regarding the rise in people accessing mental health services, suggests New Zealand clinics are also running on strict time schedules. The ORS and SRS quickly and clearly indicate what areas need to be reviewed in therapy and what areas are not as important; giving the client control of their therapy session. When the ORS is used in conjunction with the SRS, further confidence can be given in the client's therapy positively progressing, and reducing the risk of drop out, as suggested by their strong positive correlations. Together, this information highlights the need for ultra-brief outcome evaluation alternatives in New Zealand clinics and through the findings of the current study, the ORS has revealed to be a good alternative to the current HoNOS for ultra-brief outcome evaluations.

A further implication of this finding is getting therapists to religiously complete outcome evaluation scales. Gilbody et al. (2002) stated that routine outcome measuring detracts from the therapeutic relationship,

hence why clinicians chose not to do them. However, the present study's results suggest otherwise, where high ORS scores were paired with high SRS scores. Indicating that the ORS was not detracting from the therapeutic relationship, but instead supporting it. Miller et al. (2003;2006) however, suggest outcome evaluations do not get completed due to time constraints.

Outcome evaluations not only aid meaningful change, but help to build a library of raw data for future clients to be compared with and identify risks. It will likely take further education and support around outcome evaluations to engage clinicians in administering outcome measures. Ideally, time is spent with therapists during training to talk about their benefits and usefulness. However, with stretched health department budgets, the likelihood of money being spent on training for outcome measures appears small. Psychometrically adequate ultra-brief outcome measures, provide a method to change negative attitudes about outcome scales because of their overall brevity and simplicity. Miller et al. (2006) and Boswell et al. (2013) hope as time increases and new clinicians are brought in, outcome evaluations will become a routine aspect and be considered "just a part of the job", and the reluctance to complete them will fade. The present study's findings suggest this transition is slowly happening, but that more work needs to be done.

A direct practical implication of the current study is the importance and significance of rapport building between therapists and clients, to develop therapeutic alliance and to strengthen therapy outcomes. When combining the ORS and the SRS, the results suggest that a strong

therapeutic relationship increases therapy outcomes. This was shown through clients who rated their wellbeing highly, also rated their session rating higher. This is a reminder to never underestimate the impact a positive therapeutic relationship can have on changes in therapy. It also reinforces clinicians focus on developing an alliance before diving into therapy.

The current study has indicated that the ORS is a possible ultra-brief alternative to the HoNOS family of measures. It has achieved this through the ORS successfully measuring clinical change, converging with some HoNOSCA scales, and providing evidence of its psychometric properties. It is fair to say that the ORS and SRS appear appropriate for use with the New Zealand clinical population.

### **Study Limitations**

Several limitations may affect the generalisability of the findings from the current study. These include the relatively small sample size, the young average sample age, the study not having a non-clinical comparison group and the ORS not being designed in New Zealand.

The small sample size ( $N=98$ ) limits the size of effect and the results statistical power (Tabachnick & Fidell, 2007). Although the raw data set was large, unfortunately, the data was patchy and required large amounts of cleaning for it to be used for analysis. This means, that although the results may show statistical significance, the probability that effects exist is limited. It is possible that the small sample size reflects the feasibility and staff opinions about outcome measures. In future, staff need

comprehensive training to support their understanding of the benefits and usefulness of the measures before administering them.

Tabachnick and Fidell (2007) highlight that samples are measured to make generalisations about populations. The current study originally used a clinical and non-clinical population to create a fuller more wholesome sample to make generalisations about all mental health populations in New Zealand. However, after the cleaning of data was completed, it became apparent that the meaningful data was mainly collected from the iCAMHS service, which is a mental health service for children. This became a limitation for the current study because it meant the average sample age was 12 years old and not generalisable to the wider New Zealand population. The average age of a New Zealander was 38.0 years in 2013 (Statistics New Zealand, 2013), and the average non-clinical sample age was 22 years old. This also meant that the non-clinical comparison data was no longer appropriate. Apart from it being too much of a developmental difference to make fair and accurate statistical comparisons, the outcome measures take a slightly different form for different ages. Without a non-clinical comparison sample the study's findings relied on within and between groups correlational analysis. Thus, it is important for the findings of this study to not be applied as an example of all ORS information across different ages.

A further limitation is that the ORS was designed overseas without a New Zealand population in mind. This means the ORS may not specifically measure wellbeing factors that a New Zealander may view as important. It may also mean that clinical cut off scores, and significant

change measures may not be accurate with a New Zealand population. A frequent cultural limitation of the ORS discussed with stakeholders was its lack of acknowledgement to spiritual wellbeing and how the ORS scales relates to Maori health values. New Zealand's indigenous culture Maori, value spiritual wellbeing, and measuring it in practice may help the inclusiveness of therapy as well as therapy progression.

A final limitation of this study is in relation to the quantitative method chosen to collect data. Quantitative data is valid and important when making numerical comparisons, such as descriptive statistics and correlations, however it does not provide human descriptions on the qualities or characteristics of the tool (Tabachnick & Fidell, 2007). For this reason, using a mixed method approach by incorporating qualitative methods would have been beneficial. By incorporating qualitative methods, more reliable data could have been gathered regarding the feasibility of the ORS as well as HoNOS scales. Simply talking to administrators and users of the ORS and HoNOS scales would have provided a complete picture of what the numbers (or lack of) show, and what clinicians think regarding the clinical feasibility.

### **Future Directions for Research**

Three specific directions for future research have been identified. These are whether an addition of an item targeting spiritual wellbeing would add to the value and clinical significance of the ORS (with a New Zealand sample); a specific study targeting the general New Zealand population at different age groups, and examining how the ORS relates, including incorporating clinical and non-clinical samples; and lastly,

utilising a mixed method design approach where quantitative and qualitative methods are employed. Qualitative methods would allow surveys and semi-structured interviews to be used to better understand the clinical feasibility of the ORS.

Results indicated that all ethnicities in the current study had similar pre- and post- scores that all showed clinical improvement. This suggests that all ethnicities in the current study related to the ORS and SRS. However, currently there have been no studies completed on cultural differences and the ORS in New Zealand. Given that prior research has acknowledged that ethnic minorities are increasingly accessing mental health services, and have less positive therapy success rates, an incorporation of an additional scale targeted at a cultural element such as wellbeing, may provide interesting results. Future research could develop the scales to suit New Zealand's indigenous culture and accurately reflect important aspects of their wellbeing. A move towards authentically incorporating Maori into outcome evaluations, could be evaluating their spiritual wellbeing and looking at aspects of their connectedness to tapuna (ancestors), a quality often aspired to by Maori (Bennett, Flett & Babbage, 2014).

Future research would benefit from looking at the ORS more extensively within a New Zealand population. A larger sample size would provide stronger effect sizes and statistical significance with more validity. The incorporation of a non-clinical sample would allow accurate and fair comparisons to clinical data. It would be interesting for future research to break down samples into age groups and view the ORS with different age

groups to see if there is an age group it is most effective with. This was unable to be completed in the present study due to the relatively small sample size, and their being pockets of data at specific age groups.

A final future direction for extensive ORS research is to incorporate qualitative methods into the data collection. Qualitative methods can provide a fuller picture on the ORS feasibility. Future researchers would benefit from developing a survey and semi-structured interviews to evaluate whether clinicians like the scale, use the scale, and believe it is valid and reliable compared to other outcome evaluations. One of the most important elements of outcome evaluations identified in the literature, is whether clinicians actually use the scales. Gaining a first-hand understanding of why they do or do not use it, will enable the ORS to progress in New Zealand's clinical settings. The current study noticed that there were 'pockets' of data in both the ORS and HoNOSCA measures, suggesting there were staff issues or priority calls made with administering outcome evaluations. A further understanding of the collection of outcome information would help progress the industry.

## **Final Conclusions**

This research proposes that there is a need for ultra-brief therapeutic, outcome evaluations in New Zealand mental health services, and that the ORS and SRS have the potential to fill this gap. Despite the small sample size, the young average age, and the lack of a comparison group, the ORS and SRS provided evidence of measuring change, eliciting meaningful information for therapy, and being comparable to HoNOSCA. However, further analysis was prevented due to erroneous

data recording. Research has suggested some clinicians have a resistance towards administering outcome evaluations, for a range of reasons, such as the time it takes to administer, score and interpret (Miller et al. 2003; Miller et al. 2006). It is possible that clinicians from the sample population's service, found completing two outcome measures infeasible. Future research could explore the area of feasibility and outcome measures in New Zealand. Unfortunately, this erroneous data collection has spoiled possible findings of the ORS and SRS in a New Zealand population.

This study aimed to use current literature to explore the ORS and SRS to see how it responds in a New Zealand population. The findings overall suggested the ORS and SRS reflect change, relate to the current mandated outcome measure, and relate to a New Zealand population group. Ultimately, supporting the adoption of brief outcome measures. The research makes a contribution to the field of clinical psychology in New Zealand, in the everyday assessment of therapy outcomes. No prior research on the ORS and SRS among a New Zealand population was identified. This meant that any ideas or information found, may be beneficial for all clinical professionals.



## References

- American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). Washington, DC: Author.
- Ardito, R. B., & Rabellino, D. (2011). Therapeutic Alliance and Outcome of Psychotherapy: Historical Excursus, Measurements, and Prospects for Research. *Frontiers in Psychology*, 2, 270.  
<http://doi.org/10.3389/fpsyg.2011.00270>
- Barlow, D. H. (2004). Psychological treatments. *American Psychologist*, 59, 869-878.
- Beit-Hallahmi, B. (1974). Psychology of religion 1880-1930: The rise and fall of a psychological movement. *Journal of the History of the Behavioral Sciences*, 10(1), 84-90.
- Bennett, S. T., Flett, R., & Babbage, D. (2014). Culturally adapted cognitive behaviour therapy for Maori with major depression. *The Cognitive Behaviour Therapist*, 7(20), 1-16.
- Bennett, S. T., Flett, R., & Babbage, D. (2016). Considerations for culturally responsive cognitive-behavioural therapy for Maori with depression. *Journal of Pacific Rim Psychology*, 10, e8  
[doi:10.1017/prp.2016.5](https://doi.org/10.1017/prp.2016.5).
- Blow, A., Timm, T., & Cox, R. (2008). The Role of the Therapist in Therapeutic Change: Does Therapist Gender Matter? *Journal of Feminist Family Therapy*, 20(1), 66-86.
- Bohart, A. C., & Wade, A. G. (2012). The Client in Psychotherapy. Chapter 7 in Lambert.

- Bordin, E. S. (1979). The generalizability of the psychoanalytic concept of the working alliance. *Psychotherapy: Theory, research & practice*, 16(3), 252.
- Boswell, D. L., White, J. K., Sims, W. D., Harris, R. S., and Romans, J. S. C. (2013). Reliability and Validity of the Outcome Questionnaire – 45.2. *Psychology Reports: Mental and physical Health*, 112(3), 689-693.
- Bringhurst, D. L., Watson, C. W., Miller, S. D., & Duncan, B. L. (2006). The reliability and validity of the Outcome Rating Scale: A replication study of a brief clinical measure. *Journal of Brief Therapy*, 5(1), 23-30.
- Brooks, R. (2000). The reliability and validity of the Health of the Nation Outcome Scales: validation in relation to patient derived measures. *Australian and New Zealand Journal of Psychiatry*, 34, 504-511.
- Brown, J., Dreis, S., & Nace, D. (1999). What really makes a difference in psychotherapy outcome? Why does managed care want to know? In M. Hubble, B. Duncan, & S. Miller (Ed.), *The heart and soul of change* (pp. 389-406). Washington, DC: APA Press.
- Callaly, T., & Halleborne, E. L. (2001). Introducing the routine use of outcomes measurement to mental health services. *Australian Health Review*, 24, 43-50.
- Campbell, A., & Hemsley, S. (2009). Outcome Rating Scale and Session Rating Scale in psychological practice: Clinical utility of ultra-brief measures. *Clinical Psychologist*, 13(1), 1-9.

- Campbell, L. F., Norcross, J. C., Vasquez, M. J. T., & Kaslow, N. J. (2013). Recognition of Psychotherapy Effectiveness: The APA Resolution. *Psychotherapy, 50*(1), 98-101.
- Caplan, E. (1998). *Mind games: American culture and the birth of psychotherapy*. Berkeley: University of California Press.
- Chapple, S. (2000). Maori socio-economic disparity. *Political Science, 52*(2), 101-115.
- Coakes, S., & Ong, C. (2011). *SPSS version 18.0 for windows: Analysis without anguish*. Milton, Qld: John Wiley & Sons Australia.
- Coaley, K. (2009). *An introduction to psychological assessment and psychometrics*. Los Angeles: SAGE.
- Code of Ethics Review Group. (2002). *Code of ethics for psychologists working in Aotearoa/New Zealand*. NZ Psychologist Board.
- Cohen, M., Beard, C., and Bjorgvinsson, T. (2015). Examining patient characteristics as predictors of patient beliefs about treatment credibility and expectancies for treatment outcome. *Psychotherapy Integration, 25*, (2), 90-99.
- Cuijpers, P., Li, J., Hofmann, S., & Andersson, G. (2010). Self-reported versus clinician-rated symptoms of depression as outcome measures in psychotherapy research on depression: A meta-analysis. *Clinical Psychology Review, 30*(6), 768-778.
- DeSantis, B., Jackson, M. J., Duncan, B. L., & Reese, R. J. (2017). Casting a wider net in behavioral health screening in primary care: a preliminary study of the Outcome Rating Scale. *Primary health care research & development, 18*(2), 188-193.

- Eysenck, H. J. (1952). The effects of psychotherapy: an evaluation. *Journal of Consulting Psychology, 16*(5), 319-324.
- Eysenck, H. J. (2013). Learning theory and behaviour therapy. *Readings in Clinical Psychology, 349*.
- Flannery-Schroeder, E., & Lamb, A. N. (2009). Cognitive behavior therapy. In J. L. Matson, F. Andrasik, and M. L. Matson's (Eds.) Treating childhood psychopathology and developmental disabilities. (pp. 55-78). New York, NY: Springer. DOI: 10.1007/978-0-387-09530-13.
- Garralda, E., & Yates, P. (2000). HoNOSCA: Uses and Limitations. *Child Psychology and Psychiatry Review, 5*(3), 131-132.
- Gilbody, S. M., House, A. O., & Sheldon, T. A. (2002). Psychiatrists in the UK do not use outcome measures. *The British Journal of Psychiatry, 180*(2), 101-103.
- Goldfried, M. R., & Wolf, B.E. (1996). Psychotherapy practice and research: Repairing a strained alliance. *American Psychologists, 51*, 1007-1016.
- Hafkenscheid, A., Duncan, B.L., & Miller, S.D. (2010). Outcome Rating Scale (ORS) and Session Rating Scale (SRS): Psychometric findings with the Dutch translation. *Journal of Brief Therapy, 7*, 1-2.
- Howard, K., Moras, K., Martinovich, Z., & Lutz, W. (1996). Evaluation of psychotherapy: Efficacy, effectiveness, and patient progress. *American Psychologist, 51*(10), 1059-1064.
- Janse, P., Boezen-Hilberdink, L., Van Dijk, M. K., Verbraak, M. J. P. M., & Hutschemaekers, G. J. M. (2014). Measuring feedback from clients: The psychometric properties of the Dutch Outcome Rating Scale

- and Session Rating Scale. *European Journal of Psychological Assessment*, 30(2), 86–92.
- Kisely, S., Campbell, L. A., Cartwright, J., Cox, M., & Campbell, J. (2010). Do the Health of the Nation Outcome Scales measure outcome? *Canadian Journal of Psychiatry*, 55(7), 431-9.
- Lakes District Health Board. (2017). *About Lakes DHB*. Retrieved from <http://www.lakesdhb.govt.nz/Article.aspx?ID=8186>
- Lakes District Health Board. (2017). *iCAMHs Rotorua*. Retrieved from <http://www.lakesdhb.govt.nz/Article.aspx?ID=6649>
- Lambert, M. J. (1992). Psychotherapy outcome research: Implications for integrative and eclectic therapists.
- Lambert, M. J. (2001). Psychotherapy outcome and quality improvement: Introduction to the special section on patient-focused research. *Journal of Consulting and Clinical Psychology*, 69(2), 147-149.
- Lambert, M. J. (2004). *Bergin and Garfield's Handbook of Psychotherapy and Behaviour Change*. The University of Michigan: Wiley.
- Lambert, M. J. (2012). Helping clinicians to use and learn from research-based systems: The OQ-Analyst. *Psychotherapy*, 49, 109–114. doi:10.1037/a0027110.
- Lambert, M. J. (2013). Outcome in psychotherapy: The past and important advances. *Psychotherapy*, 50(1), 42-51.
- Lambert, M. J., & Bergin, A. E. (1994). The effectiveness of psychotherapy. In A. E. Bergin & S. L. Garfield (Eds.), *Handbook of psychotherapy and behaviour change* (pp. 143-189). New York: John Wiley.

- Lambert, M.J., Burlingame, G. M., Umphress, V., Hansen, N. B., Vermeersch, D. A., Clouse, G. C. (1996). The reliability and validity of the Outcome Questionnaire. *Clinical Psychology and Psychotherapy*, 3, 249-258.
- Lambert, M. J., Whipple, J. L., Smart, D. W., Vermeersch, D. A., and Nielsen, S. L. (2001). The effects of providing therapists with feedback on patient progress during psychotherapy: Are outcomes enhanced? *Psychotherapy Research*, 11, 49–68.
- Lehner, G. F. (1952). Defining Psychotherapy. *American Psychologist*, 7(9), 547. <http://dx.doi.org/10.1037/h0053853>
- Lim, W. M., & Ting, D. H. (2012). *Research methodology: a toolkit of sampling and data analysis techniques for quantitative research*. Munich: GRIN Verlag.
- Littauer, H., Sexton, H., & Wynn, R. (2005). Qualities clients wish for in their therapists. *Scandinavian Journal of Caring Sciences*, 19(1), 28-31.
- Ludy, T. M. (2007). *A Brief History of Modern Psychology*. Blackwell Publishing.
- Miller, S. D. (2012). The Outcome Rating Scale and the Session Rating Scale. *Integrating Science and Practice*, 2(2), 28-31.
- Miller, S. D., & Duncan, B. L. (2000). *The Outcome Rating Scale*. Chicago: Author.
- Miller, S. D., & Duncan, B. L. (2004). *The outcome and session rating scales: Administration and scoring manual*.

Miller, S. D., Duncan, B. L., Brown, J., Sorrell, R., & Chalk, M. (2006).

Using formal client feedback to improve retention and outcome:

Making ongoing real-time assessment feasible. *Journal of Brief*

*Therapy*, 5(1).

Miller, S. D., Duncan, B. L., Brown, J., Sparks, J., & Claud, D. (2003). The

outcome rating scale: A preliminary study of the reliability, validity,

and feasibility of a brief visual analog measure. *Journal of Brief*

*Therapy*, 2(2), 91-100.

Miller, S. D., Hubble, M. A., Chow, D. L., & Seidel, J. A. (2013). The

outcome of psychotherapy: Yesterday, today, and tomorrow.

*Psychotherapy*, 50, 88–97. doi:10.1037/a0031097

Ministry of Health. (2012). Legislation. Retrieved from

<http://www.health.govt.nz/nz-health-statistics/access-and-use/legislation>

Ministry of Health. (2012). Our History and Current Position. Retrieved

from <http://www.health.govt.nz/nz-health-statistics/access-and-use/legislation>

Ministry of Health. (2012). Mental health and addiction: Service Use

2014/2015. Retrieved from

<http://www.health.govt.nz/publication/mental-health-and-addiction-service-use-2014-15>

Ministry of Health. (2017). *Treaty of Waitangi Principles*. Retrieved from

<http://www.health.govt.nz/our-work/populations/maori-health/he-korowai-oranga/strengthening-he-korowai-oranga/treaty-waitangi-principles>

- Mitchell, M. L., & Jolley, J. M. (2013). Research design explained. Belmont, CA: Wadsworth.
- Moran, P. (2017). Selecting self-report outcome measures for use in family and systemic therapy. *Journal of Family Therapy*, 39(1), 41-56.
- Mukaka, M. (2012). A guide to appropriate use of Correlation coefficient in medical research. *Malawi Medical Journal: The Journal of Medical Association of Malawi*, 24(3), 69–71.
- Muller, M., Vandeleur, C., Weniger, G., Prinz, S., Vetter, S., & Egger, S. T. (2016). The performance of the Health of the Nation Outcome Scale as measures of clinical severity. *Psychiatry Research*, 239, 20-27.
- Munford, R., Sanders, J., Veitch, B. M., & Conder, J. (2008). Ethics and research: Searching for ethical practice in research. *Ethics and Social Welfare*, 2(1), 50-66.
- New Zealand Health Survey. (2014). Annual Update of Key Results. Retrieved from: <http://www.health.govt.nz/publication/annual-update-key-results-2014-15-new-zealand-health-survey>
- New Zealand Psychologists Board. (2011). Cultural Competencies. Wellington, New Zealand.
- Norcross, J. C. (2005). The psychotherapist's own psychotherapy: Educating and developing psychologists. *American Psychologist*, 60, 840 – 850. doi:10.1037/0003-066X.60.8.840
- Norcross, J. C. (Ed.). (2011). Psychotherapy relationships that work (2nd ed.). New York: Oxford University Press.



- Official Information Act, No. 156. (1982), New Zealand. Retrieved from <http://www.legislation.govt.nz/act/public/1982/0156/latest/DLM64785.htm?src=qs>
- Pallant, J. (2007). *SPSS survival manual: A step-by-step guide to data analysis using SPSS version 15*. Maidenhead, United Kingdom: Open University Press.
- Pirkis, J. E., Burgees, P. M., Kirk, P. K., Dodson, S., Coombs, T.J. & Williamson, M. K. (2005). A Review of The Psychometric Properties of the Health of The Nation Outcome Scales (HoNOS) Family of Measures. *Health Qual Life Outcomes*, 28(3), 76.
- Prochaska, J. O., & DiClemente, C. C. (1982). Transtheoretical therapy: Toward a more integrative model of change. *Psychotherapy: theory, research & practice*, 19(3), 276.
- Prochaska, J. O., & Norcross, J. C. (2014). *Systems of psychotherapy: a transtheoretical analysis*. Stamford: Cengage Learning.
- Rochford, T. (2004). Whare Tapa Wha: A Māori model of a unified theory of health. *The Journal of Primary Prevention*, 25(1), 41-57.
- Rogers, C. R. (1951). Studies in client-centered psychotherapy III: the case of Mrs. Oak—a research analysis. *Psychological Service Center Journal*.
- Rosenzweig, S. (1936). Some implicit common factors in diverse methods of psychotherapy. *American journal of Orthopsychiatry*, 6(3), 412.
- Shaw, S. L., & Murray, K. W. (2014). Monitoring alliance and outcome with client feedback measures. *Education*, 36(1), 43 – 57.

- Shapiro, S. D. (2002). Renewing the scientist-practitioner model. *Psychologist-Leicester* -, 15(5), 232-235.
- Smith, M. L., Glass, G. V., & Miller, T. I. (1980). The benefits of psychotherapy. Baltimore, MD: Johns Hopkins University Press.
- Stats NZ. (2013). *2013 Census Quick Stats: Rotorua Population*. Retrieved from [http://www.stats.govt.nz/Census/2013-census/profile-and-summary-reports/quickstats-about-a-place.aspx?request\\_value=13918&tabname=Culturaldiversity](http://www.stats.govt.nz/Census/2013-census/profile-and-summary-reports/quickstats-about-a-place.aspx?request_value=13918&tabname=Culturaldiversity)
- Stats NZ. (2013). *2013 Census: Latest Information*. Retrieved from <http://www.stats.govt.nz/Census/2013-census.aspx>
- Strupp, H.H., & Howard, K.I. (1992). A brief history of psychotherapy research. In D.K. Freedheim, H.J. Freudenberger, J.W. Kessler, S.B. Messer, D.R. Peterson, Strupp, H.H., & P.L. Wachtel (Eds.), *History of psychotherapy: A century of change* (pp. 309-334). Washington, DC: American Psychological Association. *A Brief Introduction to Psychotherapy Research*.
- Tabachnick, B. G., & Fidell, L. S. (2007). Using multivariate statistics (5th ed.). Boston, MA: Pearson Education.
- Taylor, M. (1999). Changing what has gone before: The enhancement of an inadequate psychology through the use of an Afrocentric-feminist perspective with African American women in therapy. *Psychotherapy: Theory, Research, Practice, Training*, 36(2), 170-17.

- Te Pou. (2012). *The HoNOS Family of Measures: A technical review of their psychometric properties*. Auckland: Te Pou o Te Whakaaro Nui.
- Te Pou. (2014). *HoNOSCA: A New Zealand clinician's guide to ratings and use*. Auckland: Te Pou o Te Whakaaro Nui.
- Trauer, T., Gill, L., Pedwell, G., & Slattery, P. (2006). Routine outcome measurement in public mental health – what do clinicians think? *Australian Health Review*, 30(2), 144-147.
- Trauer, T., & Nagel, T. (2014). Outcome measurement in adult indigenous mental health consumers. *Advances in Mental Health: Promotion, Prevention and Early Intervention*, 11(1), 1-6.
- Wampold, B. E. (2010). *The basic of psychotherapy: An introduction to theory and practice*. Washington DC: American Psychological Association.
- Wampold, B. E. (2017). *How to use research to become more effective therapists. European Family Therapy Association Series*. Washington DC: American Psychological Association.
- Wampold, B. E., Goodheart, C. D., & Levant, R. F. (2007). Clarification and elaboration on evidence-based practice in psychology.
- Wampold, B. E., & Imel, Z. E. (2015). *The great psychotherapy debate: The evidence for what makes psychotherapy work*. Routledge.
- Wilson, G.T. (1982). Psychotherapy Process and Procedure: The Behavioral Mandate. *Behavior Therapy* 13, 291–312.

## Appendices

### Appendix A. The Outcome Rating Scale

#### Outcome Rating Scale (ORS)

Name _____	Age (Yrs): _____	Sex: M / F _____
Session # _____	Date: _____	
Who is filling out this form? Please check one: Self _____ Other _____		
If other, what is your relationship to this person? _____		

Looking back over the last week (or since your last visit), including today, help us understand how you have been feeling by rating how well you have been doing in the following areas of your life, where marks to the left represent low levels and marks to the right indicate high levels. *If you are filling out this form for another person, please fill out according to how you think he or she is doing.*

**Individually:**  
(Personal well-being)

|-----|

**Interpersonally:**  
(Family, close relationships)

|-----|

**Socially:**  
(Work, school, friendships)

|-----|

**Overall:**  
(General sense of well-being)

|-----|

International Center for Clinical Excellence

[www.centerforclinicalexcellence.com](http://www.centerforclinicalexcellence.com)

© 2002, Scott D. Miller, Barry L. Duncan, & Lynn Johnson

*Note: Sample Copy. Working copies may be obtained in 25 different languages at <http://scottdmiller.com/performance-metrics/>*

## Appendix B. The Session Rating Scale

### Session Rating Scale (SRS V.3.0)

Name _____	Age (Yrs): _____
ID# _____	Sex: M / F
Session # _____	Date: _____

Please rate today's session by placing a hash mark on the line nearest to the description that best fits your experience.

#### Relationship:

I did not feel heard,  
understood, and  
respected.

|-----|

I felt heard,  
understood, and  
respected.

#### Goals and Topics:

We did not work on  
or talk about what I  
wanted to work on  
and talk about.

|-----|

We worked on and  
talked about what I  
wanted to work on  
and talk about.

#### Approach or Method:

The therapist's  
approach is not a  
good fit for me.

|-----|

The therapist's  
approach is a good fit  
for me.

#### Overall:

There was  
something missing  
in the session today.

|-----|

Overall, today's  
session was right  
for me.

International Center for Clinical Excellence

[www.centerforclinicalexcellence.com](http://www.centerforclinicalexcellence.com)

© 2002, Scott D. Miller, Barry L. Duncan, & Lynn Johnson

Note: Sample Copy. Working copies may be obtained in 25 different languages at  
<http://scottdmiller.com/performance-metrics/>