University of Strathclyde
School of Psychological Sciences and Health
Counselling Unit

The Strathclyde Inventory as a Measure of Outcome in Person-Centred Therapy

by

Susan Stephen

A Doctoral Thesis completed in order to meet the criteria for the title of Doctor of Philosophy of the University of Strathclyde.

2020

This thesis is the result of the author's original research. It has been composed by the author and has not been previously submitted for examination which has led to the award of a degree.

Signed: 

Date: 7 January 2020

Acknowledgements

I would like to take this opportunity to thank everyone who has supported me through the twists and turns of this process, in particular my research supervisors, Professor Robert Elliott and Dr Diane Dixon; my practice supervisor, Terry Daly; my partner, Colin Carson; and my dear friend on the other side of the world, Victoria Traynor. Thank you for believing in me, especially during those times when I doubted myself.

I wish to express my gratitude to the hundreds of clients and therapists who have contributed data to the archive developed by the Strathclyde Counselling and Psychotherapy Research Clinic since its inception in 2007. This thesis would not have been possible without you.

I want to acknowledge the substantial contribution made to this thesis by Laura Bell, Maha Khan, Ruth Love, Hannah MacIntosh, Melanie Martin, Rebecca Moran, Emily Price and Brigid Whitehead, the intrepid MSc students who joined my Case Study research group and enabled me to conduct the third part of this thesis. Thank you for accepting the mission and completing it with such curiosity and care.

I am grateful to PCT Scotland for sponsoring me to present my research at four national and international conferences during my period of study (see Appendix A). Thank you for investing in these opportunities for me to refine my thesis through dialogue with colleagues in the field. I hope that you will consider it money well spent; a thread within our ongoing work to develop and promote person-centred therapy in Scotland and elsewhere.

Finally, I dedicate this thesis to my parents, Valerie and Alexander Stephen. Although they did not always understand the choices that I made, they always supported me. I know that they would have been delighted to see me achieve this professional milestone and that my mother, in particular, would have made it her business to sit down and read my thesis, cover to cover.

Abstract

Person-centred therapy, like other humanistic therapies, proposes a potentiality model in which psychological growth, not simply the reduction of symptoms, is the anticipated outcome of therapy. Although substantial evidence of the effectiveness of person-centred therapy using medical model concepts exists, there is a need to develop measures that test outcome in therapy according to the person-centred theory of change.

The Strathclyde Inventory (SI) is a brief self-report instrument designed to measure congruent functioning (described elsewhere as Rogers' fully functioning person, or congruence) for use as an outcome measure in therapy.

The main purpose of this innovative three-part mixed method study was to investigate the validity of the SI as an outcome measure from multiple perspectives using data collected from a large UK-based clinical population. The first study evaluated the internal structure and reliability/precision of the instrument using the Rasch model. The second study investigated patterns of change in SI scores over the course of therapy seeking evidence of sensitivity to change, as well as convergent and construct validity. The third study tested the validity of change in SI scores as a measure of congruent functioning via a meta-synthesis of a series of eight systematic case studies examining client improvement and deterioration in therapy identified by pre-post change in SI scores.

The results supported the validity of the SI as an internally consistent and precise unidimensional instrument that is able to identify meaningful change in congruent functioning within a UK-based clinical population. A brief 12-item version of the instrument was produced. An evidence-based, theoretically coherent, developmental pathway for congruent functioning was proposed, identifying self-acceptance as a pivot point. Overall, the results of this three-part study established that change in participants' scores during therapy demonstrated a high degree of variation and proposed an explanation for different post-therapy outcomes in congruent functioning.

Table of Contents

List of Tables

List of Figures

Chapter 1: Introduction

Person-centred therapy (PCT), like other humanistic therapies, is based on a potentiality model in which psychological growth is anticipated as an outcome of therapy (Rogers, 1951, 1959, 1961/1967; Cooper, O'Hara, Schmid & Bohart, 2013; Joseph, 2017). This approach is counter to the more dominant perception within society, based on the medical model, of therapy as a means to categorise, control and eliminate symptoms of distress (e.g. Sanders, 2017; Watson, 2019). Self-report instruments that are regularly used to measure outcome in therapy are based on medical model goals, such as diagnosing and achieving symptom reduction (Levitt, Stanley, Frankel & Raina, 2005). Although there is substantial evidence of the equivalent effectiveness of PCT (e.g. Elliott, Greenberg, Watson, Timulak & Freire, 2013; Lambert, Fidalgo & Greaves, 2016), demonstrating that successful outcome in this type of therapy can be measured using concepts and goals that do not fit with the theory of the approach, Levitt et al. (2005, p.113) described this as "weighing oranges with thermometers".

Along with other contemporary PCT researchers (e.g. Patterson, 2017), Levitt et al. (2005) argued for greater development and use of measures that better fit the theoretical concept of outcome within humanistic therapies. Aligned with this call to action, there has been a recent renaissance in the development of theoretically-congruent self-report instruments in the PCT field designed to measure growth-related constructs such as *unconditional positive self-regard* (e.g. The Unconditional Positive Self-Regard Scale; Patterson & Joseph, 2006), *authenticity* (e.g. The Authenticity Scale; Wood, Linley, Maltby, Baliousis & Joseph, 2008), and the specific focus of this study, the *fully functioning person* (The Strathclyde Inventory; Freire, 2007).

In this chapter I describe the development of outcome measurement within person-centred theory and practice, explain the theoretical construct that I have used to define outcome in this thesis, introduce myself and my aims in the context of this study, then finally provide an overview of the structure of my thesis.

The Origin of Outcome Measurement in Person-Centred Theory

As indicated by my use of the word 'renaissance', measuring outcome from a theoretically-congruent perspective is not new to PCT research. Indeed, it was at the heart of the research activities that informed the development of theory and practice in PCT from its earliest days (e.g. Rogers & Dymond, 1954; Barrett-Lennard, 1998; Seeman, 2008). Rogers was dissatisfied with contemporary ideas about the desired outcome of psychotherapy, describing them as "too slippery, too relativistic, to have much value in a developing science of personality" (Rogers, 1963, p.17). As a scientist, Rogers criticised both the lack of specificity and cohesion in the concept of 'positive mental health' and also argued that diagnostic categories were not scientific concepts. As a counsellor, he opposed the idea that successful therapy merely adjusted people to society and questioned if becoming 'normal' could be considered an appropriate outcome of therapy. In contrast, Rogers had observed through his own practice that therapy resulted in personality change and growth. He described the process as being "midwife to a new personality… the emergence of a self" (Rogers, 1951, p.xi).

Rogers and his colleagues at the University of Chicago conducted a comprehensive research program into the process and outcome of PCT (at that time known as client-centred therapy). In these studies, the *self-concept* was the dependent variable, the theoretical construct that the researchers predicted would change, a process outlined by Rogers in his 19 propositions (Rogers, 1951, pp.481-533). The SIO-Q-Sort, based on Stephenson's Q-technique method (1953, cited in Butler & Haigh, 1954), was the method used to test this developing theory of personality change.

Participants were invited, before counselling and at follow up, to sort a set of one hundred cards, each offering one personal characteristic, into nine piles (ranging from "like-me" to "unlike-me") to describe themselves as they saw themselves that day, the *self-sort*, and then to re-sort the cards into nine piles (ranging from "like-ideal" to "unlike-ideal") to describe their ideal person, the person that they would most like to be, the *ideal-sort* (Butler & Haigh, 1954, p.57).

It was proposed that discrepancies in the comparative positioning of these characteristics for ideal and self, measured using correlations, was an indicator of the participant's self-esteem, and that the mean pre-post difference between self-ideal correlations was a measure of change in self-esteem over the course of therapy. Other instruments - the Thematic Apperception Test, the Self-Other Attitude Scale and the Willoughby Emotional-Maturity Scale - were also used within the program, providing triangulated results that supported the researchers' hypotheses that personality change occurred in a direction that could be defined as an improvement, including increased emotional maturity and change in both the individual's perception of self and in their attitudes towards others (Rogers & Dymond, 1954, pp.36-37).

Integrating the results of this ongoing research program into his theoretical framework, formalised by his 1959 statement, Rogers proposed a new theoretical construct, the *fully functioning person*. He described this hypothetical person as "the end-point of optimal psychotherapy… the kind of person who would emerge if therapy was maximal" (Rogers, 1963, p.18), outlining three main characteristics: openness to experience; existential living; and trust in their own organism. He stated that although he perceived this process of change to be "quite unitary" (p.18), he had deconstructed it in order to enhance the clarity of his description of the phenomenon that he had observed. He noted that, as clients developed characteristics of the fully functioning person, they could "tolerate a much wider range and variety of feelings, including feelings which were formerly anxiety-producing; and that these feelings are usefully integrated into their more flexibly organized personalities" (p.22), enabling those individuals to respond in more constructive, creative ways to difficult life experiences.

Seeman (2008), a member of Rogers' original team, continued to investigate the concept of the fully functioning person throughout his career, conducting over 30 studies to detect and measure change within the person at subsystem level while using two instruments, the Tennessee Self-Concept Scale and a peer judgment scale, to sort participants as 'high' or 'non-high' functioning persons

(2008, p.242). This work resulted in the development of his *Human Systems Health Model* (2008, p.149), a whole-person system that depicted an interwoven relationship between bi-directional vertical subsystems: interpersonal/ecological (person-to-environment, person-to-person); cognitive; pre-cognitive; perceptual; physiological; biochemical. According to Seeman, "the fully functioning person is characterized by an optimal level of organismic connectedness and integration" (2008, p.156).

Outcome as Process

Rogers made an important statement that is fundamental to understanding the concept of the fully functioning person:

> Since some of these terms sound static, as though such a person "had arrived", it should be pointed out that all the characteristics of such a person are *process* characteristics. The fully functioning person would be a person-in-process, a person continually changing. Thus, his specific behaviors cannot in any way be described in advance. The only statement that can be made is that the behaviors would be adequately adaptive to each new situation, and that the person would be continually in a process of further self-actualization (1963, p.234).

The idea of successful therapeutic outcome as the development of a way of processing, rather than a state, is a fundamental characteristic of the model of change underpinning PCT. Rogers's thinking straddled two paradigms – the predominant Cartesian-Newtonian paradigm and an emergent organismic scientific paradigm (e.g. Ellingham, 2001; Warner, 2009). This can be seen in the continuing development of concepts such as the *actualising tendency* (Rogers, 1959, p.195), re-emphasised as a verb rather that a noun – i.e. "the organism tends to actualise" (Tudor, 2008a, p.70) - with actualisation as a process in which "the organism is always in motion [… a] stream of felt processes, embodying or leading to symbolic meanings, to which the term *experiencing* [author's emphasis] is given" (Barrett-Lennard, 1998, p.76).

It was Rogers' collaboration with Gendlin – whose work expanded the concept of experiencing, the foundation for focusing-oriented therapy (Gendlin, 1996) and emotion-focused therapy (Elliott, Watson, Goldman & Greenberg, 2004), two "distinctive tribes… within 'one [person-centred] nation'" (Warner, 2000, pp.37-38) – that facilitated this transition in Rogers' work, in particular in the development of a *process conception of psychotherapy* (Barrett-Lennard, 1998, p.83). Rogers (1961/1967, p.131) wrote:

> Individuals move, I begin to see, not from a fixity or homeostasis through change to a new fixity, though such a process is indeed possible. But much the more significant continuum is from fixity to changingness, from rigid structure to flow, from stasis to process.

This was a new but clearly related conceptual model of the potential outcome of PCT, presented as a continuum of change occurring within the process of psychotherapy. For the purpose of illustrating change within this continuum, Rogers (1961/1967) described seven stages but emphasised that this number was arbitrary, noting that its form as a continuum, within which there may be any number of intermediate points, was its most important characteristic. The seventh stage corresponded with his concept of the fully functioning person in which "the client has now incorporated the quality of motion, of flow, of changingness, into every aspect of his psychological life, and this becomes its outstanding characteristic" (p.154). In contrast, he described people at the first stage as fixed and remote in their experiencing, unable to own their feelings and unwilling to communicate about themselves. Rogers noted that it was unlikely for people at this first stage of the continuum to voluntarily come into therapy (p.132).

Rogers' work to conceptualise the stages on this continuum created a descriptive framework within which a process of moving towards becoming fully functioning could begin to be observed and measured. In the process of developing a corresponding *Process Scale* (Walker, Rablen & Rogers, 1960), the concept was further elaborated into seven "threads, separable at first, becoming more of a unity as the process continues" (1961/1967, p.156): feelings and personal meanings,

experiencing, incongruence, communication of self, construing of experience, relationship to problems, and manner of relating. In Rogers' next major research program, the Wisconsin Study, four of these seven scales were extracted for the assessment of change: personal constructs (i.e. feelings and personal meanings), experiencing, problem expressions (i.e. relationship to problems), and manner of relating (Barrett-Lennard, 1998, p.271). Then, over the course of the next few years, the experiencing strand was reformulated as the *Experiencing Scale* (Klein, Mathieu-Couglan & Kiesler, 1986. p.26-27), an observer-measure of in-session client experiencing that is still in widespread use today, in particular to test the association between depth in client experiencing within therapy and outcome, usually measured using symptom-related measures (e.g. Goldman, Greenberg & Pos, 2005; Toukmanian, Jadaa & Armstrong, 2010; Pascual-Leone & Yeryomenko, 2017).

Outcome in Practice

Tudor, Keemar, Tudor, Valentine & Worrall (2004) criticised Rogers' process conception theory because of the tendency for it be used as a stage model; an argument that could also be extended to the Experiencing Scale. They contended that, while it was clear that Rogers intended it to be "descriptive rather than predictive or diagnostic" (p.46), in their experience students and practitioners often become absorbed in the distinction between stages, losing sight of Rogers' description of process as a continuum. Tolan (2017) confirmed this view, noting that in her experience it was unhelpful to locate a particular client within a particular stage. She wrote "It is the general direction that is important rather than any one signpost" (p.112).

Indeed, Tudor et al. (2004) proposed potential consequences in identifying a hypothetical end-point for optimal therapy. They suggested that there was a danger that Rogers' description of the fully functioning person may become aspirational for therapists in their client work, introducing therapist-driven goals that undermine the principle of non-directivity and compromise the effectiveness of the therapy itself. They emphasised that it was "simply not humanly possible" (p.48) for the

client to receive the consistent and optimal conditions required in the hypothesis that Rogers proposed.

Instead, person-centred practitioners have tended to privilege a principled non-directiveness (Grant, 1990) and a responsive, idiosyncratic form of practice (e.g. Keys, 2003). In this way, a focus on measuring outcome in PCT, certainly in everyday counselling settings within the UK, did not develop until the combined political pressures of the drive for *evidence-based practice* (e.g Ollendick, 2014), the countermovement of *practice-based evidence* (e.g. Stiles, Barkham, Twigg, Mellor-Clark & Cooper, 2006) and the UK government's policy of austerity, which increased competition for the funding of statutory and voluntary counselling services from already scarce resources, and promoted *routine outcome monitoring* using standardised/generic measures. These measures reflected a medical model view of outcome, the reduction of symptoms, and as a result alienated person-centred practitioners from the practice of measuring outcome. I will discuss this further in Chapter 3.

Development of the Theoretical Construct: Congruent Functioning

As perhaps has become evident, several overlapping, arguably synonymous, constructs developed within person-centred theory from its earliest days, and have been used to define the nature of change and perceived outcome of PCT. Therefore, it seems important that I identify the theoretical construct that I have investigated in this thesis.

Contemporary person-centred therapists have tended to focus in practice on the concept of *congruence* as the "definition of psychological health" (Haugh, 2001a, p.1) within person-centred theory, although some prefer the alternative term, *authenticity* (e.g. Schmid, 2001; Tudor, 2008b; Joseph, 2016). There is a strong overlap between the idea of congruence and the concept of the fully functioning person. Indeed, Haugh (2001a, p.4) noted that all of the characteristics of the fully functioning person proposed by Rogers in his 1959 paper were either definitions or outcomes of congruence.

The concept of congruence as a description of psychological health emerged from the idea of *incongruence*. Rogers (1951, p.510) proposed that psychological tension exists "when the organism denies to awareness significant sensory and visceral experiences, which consequently are not symbolized and organized into the gestalt of the self-structure". Later Rogers described this experience of psychological tension as "a state of incongruence" (1959, p.213), a "discrepancy between the actual experience of the organism and the self picture of the individual insofar as it represents that experience" (1957/1992, p.828), and hypothesised that, as a result of the type of therapeutic relationship that he outlined, the client becomes "more congruent, more open to his experience, less defensive" (1959, p.218), adding that the consequence of this increase in congruence was that "tension of all types is reduced – physiological tension, psychological tension".

In one of his earliest books, Rogers (1942) conceptualised this process as the achievement of *insight*; a gradual but spontaneous experience occurring through the release of feelings and "emotionalized attitudes" (p.216), resulting in a change in perception of relationships, new willingness to accept all aspects of self, and a recognition that a choice of goals exists.

By the time Rogers presented the *Nineteen Propositions* (1951, pp.481-533), he described a process in which "under certain conditions, involving complete absence of any threat to the self-structure, experiences which are inconsistent with it may be perceived, and examined, and the structure of self revised to assimilate and include such experiences" (1951, p.517). For him at that time, congruence was the "accurate matching of experiencing and awareness" (Rogers, 1961, p.339).

Rogers (1957/1992) introduced therapist congruence as one of six necessary and sufficient conditions for therapeutic change to take place. Perhaps as a result, focus shifted to exploration of the concept of congruence as defining the quality of therapist presence (Barrett-Lennard, 1998, p.65). Indeed, much more can be found in the person-centred literature about the experience of congruence from the perspective of the counsellor than of the client (e.g. Wyatt, 2001). In an attempt to synthesize the vast range of ideas in the original and contemporary literature

describing congruence, Cornelius-White (2007) proposed a five-dimensional model that incorporated flow, genuineness, symbolization, authenticity, and organismic integration.

In her introduction to a special issue on congruence and incongruence for the journal, *Person-Centered and Experiential Psychotherapies,* Grafanaki (2013, p.183) acknowledged that often congruence and incongruence have been represented as "personal traits or characteristics, rather than fleeting moment-by-moment states of experiencing" and argued that, in contrast, contemporary research findings have shown that these "highly fluid experiential states […] are affected by the quality of interaction within and between people in any given moment". She noted that authors writing in the special issue emphasised that the ongoing process of becoming more congruent is an experience shared by both client and therapist within the therapeutic relationship.

Building on Haugh's (2001b) discussion of the theoretical incoherence in the conceptualisation of congruence, captured by the question "how can [anyone] know at any given moment, whether they are being congruent […] as the individual cannot be aware of their incongruence because the incongruent experience is out of their awareness" (p.121), Purton (2013) proposed that, instead of being a mis-match between awareness and experience, incongruence can be understood better as an unexpressed non-linguistic response to an unexamined situation. Therefore, he argued, the experience *is* within the person's awareness *if* they were to pause and consider the details of their situation and that the process of therapy is one in which the therapist assists the client to reflect further on the "*whole of their situation* [author's emphasis]" (p.196) so that they can elaborate their understanding of their response to it and articulate this more clearly. For Kuba (2013), this process involves *recognition*: sensing and acknowledging that something is there, through sincerely attending to it and accepting responsibility for it.

Vaidya (2013) proposed a re-emphasis on congruence as a process - "adequate capacity to process specific life experiences" (p.212) - and noted that,

from this perspective, a distinction between the concepts of authenticity (an attitude) and congruence (a capacity) can be identified (p.217). Furthermore, he claimed that, as a time lag is inevitable between experiencing and the symbolisation of that experiencing in a way that can be integrated into the self-concept, then incongruence between the experiencing organism and even the most fluid symbolising self is itself inevitable, an ontological experience from which relative "degrees of experience-specific and context-bound congruence can be gained" (p.218). These views, rooted in existential philosophy, are echoed by Neville (2013) who noted that "anxiety comes from being human […] congruence involves embracing that anxiety and finding ways to live with it" (p.225), contending that "client-centered therapy works at the client's point of discomfort [… and] moving towards congruence starts with allowing this sense of discomfort to come into awareness" (p.234).

Encouraged by this focus on processing in these recent considerations of the concept of congruence, and having a desire to respect Freire's (2007) use of both terms - the fully functioning person, and congruence – in her development of the Strathclyde Inventory, I have decided to use a hybrid term to represent this theoretical construct throughout my thesis: *congruent functioning*. In my opinion, this name brings together the two familiar and synonymous terms, while emphasising the processing, not static, nature of the construct. In doing so, I am not introducing a new term but using alternative language that is present in the literature: Mearns (1997, p.94) employed the term when discussing the personal development work required of trainee person-centred counsellors. Indeed, the term has been adopted as a definition of "wellness" beyond the person-centred literature. I discovered through an online search that the term is featured on flashcards (e.g. "integrated, congruent functioning aimed towards reaching one's highest potential" and "a measure of optimal health, an expression of integrated or congruent functioning") used in the training of gerontologist nurses, among others (e.g. Brainscape, n.d.).

The Researcher

I embarked on this study as the result of a range of ethical, pragmatic and theoretical considerations. At that time, I was coordinator of the Strathclyde Psychotherapy and Counselling Research Clinic ('the research clinic') and therefore had access to a substantial archive of data collected since 2007. Ethically, I felt a responsibility to ensure that the data did not remain dormant, gathering dust in filing cabinets, but was used as comprehensively as possible in accordance with the trust placed in us by the clients who participated in the research clinic's research activities and provided us with their data in good faith.

Given the wide-ranging nature of the data collected, many potential studies would have been possible, as Elliott (2009) illustrated in his blog entry *Twenty-Five Fun Research Projects to Do in a Therapy Research Clinic*. So, what would be the study that felt like the right commitment for me? I recognised quickly that ten years' worth of data collected using the Strathclyde Inventory (Freire, 2007) in the research clinic from clients participating in PCT, data that was largely undisturbed, offered me a unique opportunity to not only advance the development of the instrument itself but also to test theory at the heart of PCT. According to their scores on the instrument, do clients leave therapy changed in the way that theory predicts? This was an extremely exciting prospect for me that integrated my passion and curiosity as a person-centred therapist and researcher, and also as an advocate for the person-centred approach.

With this enthusiasm came a recognition of my intrinsic bias: I *wanted* the data to demonstrate that the theory was credible. As a result, it has been my intention throughout to let the data lead my way; to be as transparent as possible about my decision-making regarding the design of my studies and the reporting of my analyses and results and, equally importantly, to maintain an awareness of the limitations inherent in the methodology and context of this study.

Aims of the Study

The main purpose of this three-part mixed method study was to investigate the validity of the Strathclyde Inventory as a measure of outcome in PCT using the data collected from clients of the research clinic since 2007. In doing so, I was guided by these two overarching questions:

1. Are scores on the Strathclyde Inventory a valid measure of congruent functioning when used with a UK-based clinical population?
2. Do scores on the Strathclyde Inventory change over the course of therapy in a way that is consistent with person-centred theory?

In other words, I intended to test the credibility of the Strathclyde Inventory and, if demonstrated, communicate my results as a basis on which to: (a) increase the instrument's visibility, (b) encourage its use by other person-centred therapists and researchers who are interested in measuring the outcome of their work in a theoretically-congruent way, and (c) influence societal ideas about what the outcome of therapy can look like.

Structure of the Thesis

My thesis is structured in seven chapters. This introduction has presented an overview of the effort to define and measure outcome in the development of person-centred theory and practice and explained my use of the term congruent functioning within this thesis. In Chapter Two, I review a range of self-report instruments that have been developed to measure theoretical constructs synonymous or closely related to congruent functioning, and introduce the Strathclyde Inventory. In Chapter Three, I explain the importance and potential contribution offered by measurement in counselling and also the inherent methodological and ethical challenges, linking this to the opposing epistemological positions of the two disciplines, and outline the approach that I have taken in this research. In Chapters Four to Six, I present three empirical studies: the first study in which I evaluated the internal structure and precision/reliability of the instrument using the Rasch model; the second study in which I investigated patterns of change

in SI scores over the course of therapy seeking evidence of sensitivity to change, as well as convergent and construct validity; and the third study in which I tested the validity of change in SI scores as a measure of congruent functioning via a meta-synthesis of a series of eight systematic case studies examining client improvement and deterioration in therapy identified by pre-post change in SI scores. Finally, in Chapter 7, I discuss the implications of my findings for theory, research and practice and identify the contribution to knowledge that is offered.

Chapter 2: Outcome Measures in the Person-Centred Literature

Chapter Overview

As outlined in the Introduction, the development of the Strathclyde Inventory has occurred within a renaissance of interest in measuring psychological growth in therapy. In the first half of this chapter I review a range of self-report instruments that have been developed to measure a range of constructs synonymous or broadly associated with the type of psychological growth - or congruent functioning – predicted by person-centred theory. Some of these constructs aim to capture the *whole* experience (the fully functioning person, self-actualisation, self-determination, and authenticity); others focus on *specific aspects or processes* within the experience of congruent functioning (self-discrepancy, unconditional positive self-regard, self-compassion, and emotion regulation).

This is not a systematic review of the literature. This is due, first, to the potential range and variation of search terms and, second, the likelihood of becoming submerged in the extensive literature on self-esteem, emotion regulation, self-determination theory, eudaemonic / psychological well-being and positive psychology in general. Instead I have focused on investigating self-report instruments that have been identified in the person-centred literature and that have been, or could be, used for outcome measurement according to person-centred theory. In doing so, I have collated and extended similar reviews (Watson & Watson, 2010; Patterson, 2017; Zech, Brison, Elliott, Rodgers & Cornelius-White, 2018) as I made links through my reading to other related instruments. Some of the instruments or studies that I present have been developed by researchers who are either not aligned to the person-centred approach or have written their reports for a mainstream psychology readership. Therefore, there is occasional reference to the use of these instruments for activities such as diagnosis that are not associated with PCT. I have included these findings in my review if they seem relevant to an understanding of the instrument for use with individuals in clinical practice.

Having introduced the existing field of instruments, in the second half of this chapter I introduce the Strathclyde Inventory (Freire, 2007), the instrument under investigation in this thesis, and outline its development.

Overview of Existing Measures: Capturing the Whole Experience

In this section I present self-report instruments designed to measure constructs that are synonymous or broadly associated with congruent functioning: the fully functioning person, self-actualisation, self-determination and authenticity.

The Fully Functioning Person

As noted in Chapter 1, Seeman (2008) used the *Tennessee Self-Concept Scale* (TSCS; Fitts, 1965, cited in Tzeng, Maxey, Fortier & Landis, 1985) in his extensive program of study of the fully functioning person, primarily as a means of classifying participants as 'high' or 'non-high' functioning persons before testing and comparing other attributes.

According to Tzeng et al. (1985), the TSCS was developed to measure a theoretical construct of the self-concept based on three internal (*identity, self-satisfaction and behaviour*) and five external (*physical self, moral-ethical self, personal self, family self, and social self*) reference points that influenced how a person viewed themselves. The TSCS comprised 90 items: three positive and three negative items for each of fifteen intersecting categories produced by a two-dimensional grid formed from these two groups of referents (Tzeng et al., 1985). Despite a lack of research into its internal structure and sub-scales (Bentler, 1972, cited in Tzeng et al., 1985), the TSCS became "popular" (Tzeng et al, 1985, p.64) and "widely used in clinical and research settings" (Bishop, Walling & Walker, 1997). More recently, concern about its construct validity has emerged. Conducting an analysis of data collected from one clinical sample (N=132) and two college groups (N=264), Tzeng et al. (1985) were unable to find more than four distinguishable factors. In a replication of this study, focused on a more homogenous sample of participants, Bishop et al. (1997) reported similar findings. They identified evidence of multidimensionality within the sub-scales and factors that did not conform to the

proposed theoretical model. Indeed, an exploratory factor analysis of their data extracted a single general dimension. Nevertheless, the TSCS continues to be used in studies measuring change in self-concept over the course of therapy (e.g. Toukmanian at al., 2010).

Cartwright and colleagues (Cartwright & Mori, 1988; Cartwright, deBruin & Berg, 1991) designed the *Feelings, Reactions and Beliefs Survey* (FRBS) to measure nine variables drawn from Rogers' theory of personality: *focusing conscious awareness, openness to feelings in relationships, trust in self as organism, feeling uncomfortable with people, struggling with feelings of inferiority, feeling ambivalent in relationships, openness to transcendent experiences, religio-spiritual beliefs*, and *the fully functioning person* (FFP). Testing the instrument with large samples of college students, Cartwright et al. (1991) reported results relating to internal consistency, convergent and discriminant validity that, they proposed, supported the use of the FRBS in research studies of group data but not for the interpretation of individual scores. The final version of the FRBS contained 130 items, of which 15 items comprised the FFP sub-scale (e.g. I know my self-esteem is…; I often tell people my life is…). The researchers found this sub-scale to have high positive correlations with three scales (*time-competent, self-regard, self-acceptance*) of the Personal Orientation Inventory (see below) and discovered a significant negative correlation with anxiety, confirming that "persons who score high on FFP report deep feelings of security, have high self-esteem, accept themselves, and enjoy living in the here and now." (p.152-153).

The FRBS has been used in studies focused on various aspects of personality, for example, multidimensional perfectionism (Ashby, Rahotep & Martin, 2005), enviousness (Gold, 1996), and purposefulness (Cartwright & Peckar, 1993), and for the validation of other related instruments (e.g. Watson & Lilova, 2009). A German language version was created, validated and tested by Höger (1995, cited in Eckert, Höger & Schwab, 2003, p.7) to compare scores collected from a clinical and a 'normal' sample. Significant differences were found between the two groups, which were particularly large in four of the nine scales, including *fully functioning person*.

More recently, Proctor, Tweed & Morris (2016) conducted two studies to consider the concept of the fully functioning person from a positive psychology perspective. Proctor et al. assembled a battery of eight recently developed instruments that evaluated characteristics that they associated with the concept of the fully functioning person (e.g. the *Satisfaction with Life Scale, Short Depression-Happiness Scale; Authenticity Scale*, and *Organismic Valuing Scale*) and collected data online from 337 individuals aged 16 to 19 years. Using both 'model fit' and 'variance accounted for' analytic perspectives, Proctor et al. identified a single factor in their data that, in their view, reflected the concept of the fully functioning person. They did not promote the development of a new instrument based on this finding and recommended caution in drawing conclusions from their results.

Self-actualisation

The *Personal Orientation Inventory* (POI; Shostrom, 1964) was originally conceived as a therapeutic tool, "a diagnostic instrument that gives the new patient a measure of his current level of positive health or self-actualization… a 'launching pad' for the process of therapy which would suggest directions for growth towards health" (p.207). Participants were presented with 150 paired items and asked to choose the statement in each pair that most closely reflected their approach to life. These paired items were scored twice: first, according to two ratio scales, *time competence/incompetence* and *inner (towards self)/outer (towards others) orientation*; then according to ten sub-scales representing five themes: *valuing, feeling, self-perception, awareness*, and *interpersonal sensitivity* (Shostrom, 1966, cited in Tosi & Lindamood, 1975)

Shostrom (1964) tested the POI with a range of clinical and non-clinical samples. The results demonstrated that scores on the POI did distinguish between self-actualised, 'normal' and non-self-actualised groups on all but one sub-scale. Shostrom noted that the scores of clients new to therapy tended to indicate other-directedness, whereas those in their first or second year of therapy tended to score highly on inner-directedness, while those later in their therapy experience were more balanced in their inner and other orientation (pp.210-211). He reported that

the items that seemed to differentiate most clearly were those relating to social pressures, expectations and goals (p.212) and suggested that self-actualised people did not feel the need to conform, lived more fully in the here-and-now, were autonomous yet sensitive to other people, and were no longer concerned with the dichotomy between inner and other directedness, having become expanded and integrated.

Tosi and Lindimood (1975) reviewed the range of psychometric evidence generated about the POI, suggesting that overlap in the subscales resulted in an "apparent lack of parsimony" (p.221). They recommended that the POI could be used in therapy as a "stimulus for self-awareness and self-exploration" (p.222) but discouraged its use as a diagnostic instrument without further investigation. Weiss (1987) went further in his criticisms. He summarized concerns about the internal consistency of the instrument, resulting in a shifting definition of the self-actualisation construct, and concluded that should not be used in clinical or research work in its present form.

Later, Jones and Crandall (1986) argued that the POI and the related *Personal Orientation Dimensions* (POD; Shostrom 1975, cited in Knapp & Knapp, 1978), although supported by "extensive validation [and…] a tremendous amount of research" (p.64), were too long and therefore impractical. They developed a 15-item short index, the *Self-Actualisation Scale* (SAS), based on modified items drawn from the POI and POD and to be used with a four-category scale ranging from 'agree' to 'disagree'. A principal components analysis with varimax rotation indicated five factors: *autonomy or self-direction, self-acceptance and self-esteem, acceptance of emotions and freedom of expression of emotions, trust and responsibility in interpersonal relations*, and *ability to deal with, rather than avoid, undesirable aspects of life*. However, further examination of the SAS, for example by Faraci and Cannistraci (2015), identified inadequacies in the factorial structure proposed by Jones and Crandall (1986), perhaps the result of cross-cultural differences, and recommended caution when using the SAS.

Self-determination

Self-determination theory (SDT), one of the major theories informing the positive psychology movement, shares many fundamental assumptions with person-centred theory (Patterson, 2017). According to Deci and Ryan (2000), SDT proposes that humans are driven by three innate psychological needs: competence, autonomy and relatedness. Sheldon and Deci (1993, cited in Sheldon, 1995) developed the *Self-Determination Scale* (SDS), in order to "capture the prototype of the grounded and self-determined person" (Sheldon, 1995, p.28). The SDS contained ten items, representing two factors: *self-contact* and *choicefulness*. Each item consisted of two statements and participants were asked to rate which of the two statements feels most true to them using a 9-point scale ranging from 'only A feels true' to 'only B feels true'. Scores on the SDS correlated with measures of creativity, autonomy, and striving towards self-determination (Sheldon, 1995), and daily well-being (Sheldon, Ryan & Reis, 1996).

Authenticity

The concept of authenticity has fascinated scholars for thousands of years, from the writings of Aristotle to the current wave of positive psychologists (Joseph, 2016). Rogers considered congruence and authenticity to be synonymous terms (Tudor, 2008b, p.167).

Goldman and Kernis (2002) developed the *Authenticity Inventory* (AI), a 44-item instrument formed by four scales that reflected a multicomponent conceptualization of authenticity: *awareness, unbiased processing, behaviour*, and *relational orientation*. They viewed these four components as "related to, but separable from, each other" (p.19). Testing of the instrument with data collected from psychology students indicated that greater authenticity was related to higher levels of self-esteem and life satisfaction, providing evidence for the validity of the instrument.

Bond, Strauss and Wickham (2018) proposed the development of a short 20-item version of the AI, which retained the four component sub-scales and maintained a high level of internal consistency. Correlational analyses indicated that

individuals with greater authenticity experienced low attachment avoidance and anxiety, were more task-oriented and friendly, less emotionally reactive, and may have “a slight tendency towards unconventional thinking” (p.5). Their results also suggested that authenticity is not simply about feeling good and may lead the individual to challenge uncomfortable aspects of themselves. Bond et al. noted that the subscales struggled to differentiate individuals at the lower and higher ends but proposed that this was not a concern as the instrument was only used for research not diagnosis (p.5).

Wood et al. (2008) proposed a three-part definition of authenticity based on person-centred theory: *self-alienation*, *authentic living*, and *accepting external influence*. Based on data collected from 200 undergraduate students (study 1), an ethnically diverse community sample (n=180) and a separate sample of 158 undergraduate students (both samples used in study 2), Wood et al. developed and tested the 12-item *Authenticity Scale* (AS). Study 1 provided preliminary evidence for the validity of the scale, comprising three factors that fit the proposed model, in which each subscale correlated with measures of happiness, and of anxiety and stress. The structure of the scale was tested further in Study 2, which presented a comparison of fit between a 3-factor and 1-factor model as well as factor variance across samples, gender and ethnic groupings. The results demonstrated the robustness of the 3-factor model, that the AS behaved consistently across time and diverse demographic groups, and that scores correlated with measures of self-esteem, subjective well-being and psychological well-being.

Since this initial publication, the AS has been translated into many languages including French (Grégoire, Baron, Ménard, & Lachance, 2014) and Serbian (Grijak, 2017) and used in multiple studies investigating the association between authenticity and well-being, including happiness, health, and sense of self (Joseph, 2016). The AS has also been used as a measure of authenticity in studies exploring its role as a mediator between self-focused attention and well-being (Boyraz & Kuhl, 2015), and between attachment style and affective functioning (Stevens, 2017).

Highlighting reservations about using either the AI or the AS in the context of organisational research, Knoll, Meyer, Kroemer and Schröder-Abé (2015) decided to create a hybrid instrument that drew from both. They conducted a series of analyses involving five samples drawn from university settings, in which they began with 57 items drawn from the AI and the AS, and produced a new German-language 8-item scale representing a two-dimensional model of authenticity (*authentic self-awareness* and *authentic self-expression*) for use in research into the role of authenticity within organisations.

Overview of Existing Measures: Focusing on Specific Aspects or Processes

In the next section, I present self-report instruments that focus on specific aspects or processes that contribute to the development of congruent functioning according to person-centred theory: self-discrepancy, unconditional positive self-regard, self-compassion, and emotion regulation.

Self-discrepancy

The roots of self-discrepancy theory can be seen both in Rogers' personality theory (1959), in which he hypothesed that improved psychological functioning results from developing a closer fit between self-concept and organismic experience, and in the use of the Q-sort method within the research conducted by him and his team at the University of Chicago (Rogers & Dymond, 1954). Higgins, Klein and Strauman (1986) developed the *Selves Questionnaire* (SQ) to test Higgins' model of self-discrepancy. The SQ required participants to list up to ten traits or attributes that they associated with themselves from three self-perspectives (actual, ideal and ought) and from four different standpoints (own, father, mother, closest friend): twelve lists in total. Next, participants were asked to use a 4-point scale to rate the extent to which the opinion of each of these four standpoints (including self) about their actual, ideal and ought self was relevant or meaningful to them. The data was scored, first, by comparing pairs of self-concepts (e.g. actual/own and ought/own) to identify matches (synonyms) and mismatches (antonyms), then, by subtracting the number of matches from the total number of mismatches. Therefore, the potential self-discrepancy score could range from +10

to -10. Only the data collected from one 'other' standpoint was scored: the person whose opinion was most relevant to the participant. Finally, a total self-concept discrepancy score was calculated by combining the scores of four discrepancies: *actual/own-ideal/own, actual/own-ought/own, actual/own-ideal/other,* and *actual/own-ought/other*. This score was used in correlational analyses with scores obtained on a variety of other instruments to test the researchers' hypotheses. In this early study, Higgins et al. found associations between actual-ideal discrepancy and dejection-related emotions and symptoms, and between actual-ought discrepancy and agitation-related emotions and symptoms.

The way in which the SQ was administered was developed in subsequent studies, for example, by not including one or all of the 'other' domains or standpoints (e.g. Higgins, Bond, Klein & Strauman,1986; Veale, Kinderman, Riley & Lambrou, 2003) or by proposing new self-states (e.g. Cornette, Strauman, Abramson & Busch, 2009). Strauman and Higgins (1987, cited in Higgins, 1987) introduced a 4-point rating scale so that participants could indicate the extent to which the standpoint person (self or other) believed they possessed, or ought to possess, or wanted them to possess, each attribute that they listed. This enabled researchers to distinguish between synonymous matches (ratings that varied by one point) and synonymous mismatches (ratings that varied by two or more points), demonstrating the degree of mismatch.

Criticisms of the SQ included its specificity, with claims that it measured a more generalized form of self-discrepancy than proposed by self-discrepancy theory (e.g. Tangney, Niedenthal, Covert & Barlow, 1998), and the administrative burden that its methods for generating attributes and scoring placed on participants and researchers (e.g. Hardin & Lakin, 2009).

As an alternative, Watson, Bryan and Thrash (2010) introduced an online set of three instruments designed to measure self-discrepancy. The *Self-Concept Questionnaire – Personal Constructs* (SCQ-PC) is an idiographic instrument in which the participant is asked to propose bipolar constructs that describe their real, ideal and ought selves and evaluate how often each characteristic is true of each self-

construct. The *Self-Concept Questionnaire – Conventional Constructs* (SCQ-CC) is a non-idiographic instrument containing a set list of characteristics that invites the participant to identify how often each characteristic is true of their real self, ideal self and ought self. Finally, the *Abstract Measures* (AM) asks the participant to indicate their perception of the size of the discrepancies between their real and ideal selves, and their real and ought selves. For the first two instruments, a real-ideal score and a real-ought score is calculated. Watson et al. tested and compared the psychometric properties of the three instruments with clinical and non-clinical samples. They predicted that the SCQ-PC would have the strongest evidence of validity of the three instruments because it was idiographic and contained multiple items. This hypothesis was confirmed by their findings.

Watson and his colleagues proceeded to use these three instruments to test the long-term stability of real-ideal and real-ought self-discrepancies and their association with anxiety and depression in undergraduates over periods of one and three years (Watson, Bryan & Thrash, 2016) and change in these self-discrepancies, anxiety and depression between the beginning and end of therapy (Watson, Bryan & Thrash, 2014). Watson et al. (2016) found stability in scores for the two self-discrepancy constructs, anxiety and depression among undergraduates over these extended periods of time and argued that their finding of high correlations between both forms of self-discrepancy, anxiety and depression provided empirical evidence that supported Rogers' (1959) personality theory. On the other hand, Watson et al. (2014) found significant decreases in scores representing the two self-discrepancies that involved movement of the real self towards the ideal and ought selves, and lesser, but still significant, movement of the ideal and ought selves towards the real self. These changes were significantly associated with each other and with change in scores on the instruments measuring anxiety and depression. As a result, Watson et al. (2014) recommended future research on self-discrepancy as a change mechanism in therapy.

Meanwhile, Hardin and Lakin (2009) assessed the validity of the *Integrated Self-Discrepancy Index* (ISDI), designed to minimize the burden of the SQ but

retaining its idiographic format. They asked undergraduate students in two separate studies, to use a list of 100 adjectives to create five-item lists of characteristics that described their ideal and ought selves from their own standpoint and that of a significant other, then rate the extent to which these ideal and ought characteristics described their actual selves. These scores were averaged to calculate a score that represented ideal-own and ought-own discrepancies. The results provided evidence that supported the validity of the ISDI as a measure of self-discrepancy with this participant group. The ISDI has been used by researchers exploring the relationship between self-discrepancy, narcissism and self-esteem (Barnett & Womack, 2015) and between self-discrepancy and a range of emotional states (Barnett, Moore & Harp, 2017).

Nevertheless, the search for an acceptable measure of self-discrepancy continues. For example, Philippot, Dethier, Baeyens and Bouvard (2018) claimed that the ISDI fails to answer two questions: first, it does not test how closely the averaged discrepancy score fits with the participant's overall sense of that discrepancy; second, it does not investigate any relationship between the perceived self-discrepancy and any distress it may cause. Philippot et al. argued that the degree of distress may depend on the individual's attitude towards the self-discrepancy: i.e. less distress if they accept the discrepancy in self. As a result, they developed the *Self-Discrepancies Scale* (SDS) in which they aimed to bring together the best qualities of Watson et al. (2010) and Harkin and Lakin's (2009) work and produce an instrument that they believed would be useful and easy to administer in clinical settings. The SDS provided a list of 105 possible characteristics associated with competence, likeability and personal appearance, and asked participants to create two eight-item lists of desirable and negative characteristics that described their ideal self, estimate the extent to which they possessed those characteristics, and indicate the overall gap and degree of distress caused by the discrepancy; and then to do the same in relation to their "socially prescribed self" (p.71). The SDS was administered to community samples (N=218) and a clinical sample (N=60). The researchers found a high correlation between the two alternative methods for

measuring self-discrepancy. The results demonstrated that participants in the clinical sample experienced larger self-discrepancies than those in the community samples and, replicating more recent studies, found a link between ideal-self discrepancy and depression but not an association between socially prescribed self-discrepancy and anxiety, in contrast to earlier research.

Unconditional Positive Self-regard

Bozarth (2001a, p.185) proposed that unconditional positive self-regard (UPSR) fosters congruence; self-discrepancy theory offers an explanation of the mechanism through which this may occur. Patterson and Joseph (2006) developed the *Unconditional Positive Self-Regard Scale* (UPSR) by modifying items from two subscales of the Barrett-Lennard Relationship Inventory (level of positive regard, and unconditionality of positive regard). A pilot 20-item version of the scale was investigated using data collected from 210 university students. After principal components analysis, 12 items were retained, representing two independent factors: *self-regard* and *conditionality*. The results supported Patterson and Joseph's expectations regarding the validity of the sub-scales. They noted in particular a strong association with self-esteem and moderate correlations with overall mental health, proposing that the UPSR could be a useful measure for therapists wishing to evaluate their client work from a non-medicalised perspective. Griffiths and Griffiths (2013) confirmed these results and also identified a moderate negative association between UPSR and depression and anxiety. They recorded some reservations about the *conditionality* sub-scale, noting that its relatively low Cronbach alpha suggested that the sub-scale "showed questionable internal validity" (p.172), and recommended that the UPSR be scored as a single measure.

Murphy, Joseph, Demetriou and Karami-Mofrad (2017) conducted a confirmatory factor analysis of the UPSR using data collected from 230 postgraduate students. They tested 1-factor, related 2-factor and unrelated 2-factor models. Using two alternative methods, their analyses demonstrated that a related 2-factor model (indicating a relationship between *self-regard* and *conditionality*) was the best fit.

The UPSR has been used in to establish evidence of a relationship between unconditional positive self-regard and post-traumatic growth (Flanagan, Patterson, Hume & Joseph, 2015; Murphy, Demetriou & Joseph, 2015), with intrinsic aspirations (Murphy et al., 2015; Murphy et al., 2017), and authenticity (Murphy et al., 2017).

Self-compassion

Griffiths and Griffiths (2013) detected a strong and significant correlation between unconditional positive self-regard and self-compassion, as measured using Neff's (2003) *Self-Compassion Scale* (SCS). Theoretically this association makes sense. Neff (2003) proposed that self-compassion involved three aspects, self-kindness, common humanity, and mindfulness, and noted that it "can be viewed as a useful emotional regulation strategy, in which painful or distressing feelings are not avoided but are instead held in awareness with kindness, understanding, and a sense of shared humanity" (p.225). A 71-item pilot version of the SCS was prepared, which contained items selected to represent positive and negative aspects of each of the three components. A group of 391 undergraduate students were asked to rate how often they acted in the manner described by each item, using a 5-point rating scale ranging from 'almost always' to 'almost never'. Following analysis of the data collected, a 26-item version was produced, demonstrating an adequate fit to a six-factor model: *self-kindness, self-judgment, common humanity, isolation, mindfulness*, and *over-identification*.

Neff (2003) had expected a three-factor model and noted that it was likely that this more differentiated model reflected the positive and negative wording used. A single higher order factor of self-compassion was also tested (using reverse scoring for self-judgment, isolation and over-identification items) and found to fit the data "marginally well" (p.232). Expectations were met regarding convergent and discriminant validity with other instruments, in particular that self-compassion can be distinguished from self-esteem via differing associations with self-aggrandizement (p.241). The finding of a negative association with anxiety and depression, and a positive correlation with life satisfaction, supported Neff's (2003)

proposal that “self-compassion may be an adaptive process that increases psychological resiliency and well-being” (p.235).

There has been an ongoing debate in the literature about the preferred model for the SCS, as Neff, Whittaker and Karl (2017) outlined. Comparing data collected from four distinct samples, Neff et al. used confirmatory factor analyses to examine five different models: one-factor, two-factor correlated, six-factor correlated, higher order, and bifactor. They found that the six-factor correlated model provided the best fit to the data across all four samples and proposed that this may indicate “a ‘system’ view of self-compassion” (p.604). However, they noted that the bifactor model (a general self-compassion factor in addition to the six sub-scale factors) was also highly credible, accounting for 94% of the variance in total scores. Indeed, while noting that a two-factor model (i.e. *self-compassion* and *self-criticism*) may reflect the sympathetic and parasympathetic nervous systems, they argued that these two systems are in constant dynamic interaction and therefore that a single total score (representing a bifactor model) was suitable for assessing the equilibrium between the two systems.

Although for some time proposed as a change process in psychotherapy, Galili-Weinstock et al. (2018) reported the first study that examined the association between change in self-compassion and outcome on a session-by-session basis in a naturalistic psychotherapeutic setting. The participants were 112 adults attending open-ended (although limited by the academic year) psychodynamic psychotherapy at a university outpatient clinic. The full 26-item SCS was completed by at the beginning and end of therapy, while three items, each representing one of the self-kindness, common humanity, and mindfulness sub-scales, that had previously demonstrated a strong correlation with total SCS scores were extracted and completed by participants before each therapy session. A battery of other symptom-related, emotion regulation, and alliance measures were administered on either a pre-post or sessional basis. The results indicated that an increase in self-compassion levels across the therapeutic process was associated with lower levels of depression, other symptom-related and emotional difficulties at post-therapy.

The study also demonstrated that session-to-session increases in self-compassion predicted decreases in symptoms at the next session. Galili-Weinstock et al. acknowledged that asking participants to consider how self-compassionate they had been on a weekly basis may in itself have influenced their level of self-compassion and indeed the focus of their therapy, and recommended that the validity of the 3-item brief adaptation of the SCS required further investigation.

Smith, Guzman and Erickson (2018) argued that, while the SCS recognized self-compassion as a form of protection, the instrument was not designed to measure if individuals use self-compassion to protect themselves when their sense of self is under threat. They also noted concerns about the structure and length of the SCS. In response, they developed the *Unconditional Self-Kindness Scale* (USKS), stating that self-kindness can be considered a specific behavioural response to threat that involves acting in the most helpful way towards oneself in that moment. Following a literature review, Smith et al. identified the three most common threats to self: *criticism or rejection*, *failure or making a mistake*, and *becoming aware of personal flaws or imperfections*. Each threat was adapted into two questions; one that focused on an active self-kindness response (e.g. how much are you loving and kind to yourself when…?) and another that focused on a more passive self-kindness response (e.g. how much are you patient and tolerant with yourself when…?). Participants were asked to rate these six questions using a 7-point scale ranging from 'not at all' to 'a great deal'. A total of 1452 undergraduate students took part in Smith et al.'s study, completing the USKS along with a range of instruments designed to measure mental health, psychological well-being, protective factors (e.g. grit, hope, resilience), and risk factors (e.g. loneliness, perceived stress, social conflict). Smith et al reported that scores on the USKS demonstrated that it was an internally and temporally consistent scale with adequate fit to a unidimensional model. Their expectations about the associations between self-kindness, as measured by the USKS, and the other measures selected, were upheld by their results.

Emotion Regulation

Self-compassion as an "emotion regulation strategy" (Neff, 2003, p.225) makes sense when the ability to regulate emotion is understood as an aspect of congruent functioning. Behr and Becker (2002) introduced the *Scales of Experiencing Emotion* (SEE) as a short self-report measure based on constructs arising from person-centred theory, in particular symbolisation and experiencing, and from emotional intelligence; they intended the SEE to be a multidimensional instrument for use as an outcome measure in therapy. Using data collected from 456 students and members of the public, Behr and Becker (2002) reduced an original 106-item version to 46 items, which they found reflected a 7-factor solution that explained 53% of variance: *symbolization by bodily experiences, experiencing overwhelming emotions, symbolization by imagination, lack of self-control, experiencing congruence, experiencing lack of emotions,* and *regulation of emotions*. All scales were found to be internally consistent and independent of each other. Their results showed that individuals whose scores indicated that they were out of contact with their feelings were also unhappy and unclear about their feelings, and scored highly on measures of neuroticism, stress, anxiety and depression.

Next, Behr and Becker (2012) added nine items to create a 55-item version and replaced the original rating scale with a 5-point scale ranging from 'disagreement' to 'strong agreement'. They conducted further analyses of the SEE using data collected from a larger sample of students and members of the public (N= 1215), resulting in a 42-item version, then validated this version using data collected from 67 psychotherapy clients. The results of these studies replicated the original study in finding a 7-factor model, with revised names for all factors including: *accepting one's own emotions* (formerly*, experiencing congruence*) and *experiencing self-control* (formerly*, lack of self-control*). Behr and Becker (2012) noted that scores for *accepting one's own emotions* correlated positively with scores designed to measure awareness and clarity of emotions, and negatively with

scores indicating existential orientation and emotionality, proposing that accepting one's own emotions may provide a moderating effect (p.298).

The 42-item SEE was tested with a North American community sample (N=155) to investigate its cross-cultural reliability and validity (Watson & Lilova, 2009). The scores in this study replicated the 7-factor model and confirmed the SEE's overall reliability, internal consistency and validity, while noting some interesting differences between samples: in particular, the North Americans indicated that they used their imaginations less when trying to understand their emotions, had less need to hide or control their emotions, and were less aware of their emotions.

Section Summary

In this section I presented the main self-report instruments identified in the person-centred literature that have been designed to measure a range of constructs that are associated with congruent functioning. I discovered that earlier instruments have been modified or replaced over time because of various concerns raised about validity and utility in clinical practice, and that the majority of research has been confined to cross-sectional and experimental personality studies that explore and extend understanding about the relationships between the construct and other attributes. Nevertheless, I was able to locate a small number of studies that demonstrated the use of more recently developed instruments as outcome measures in clinical settings: for example, the three Self-Concept Questionnaires (Watson, Bryan & Thrash, 2014), and the Self-Compassion Scale (Galili-Weinstock et al., 2018). As a result of these studies, decades after Rogers' original research, new empirical evidence is being collected that demonstrates that psychological growth associated with congruent functioning, whether approached from the perspectives of self-discrepancy or self-compassion, is a measurable outcome of the therapeutic process.

The Strathclyde Inventory

Creation of the Strathclyde Inventory

In parallel with these recent developments, Freire (2007) aimed to develop a brief self-report outcome measure based on Rogers' concept of the *fully functioning person* (Rogers, 1961/1967, 1963, 1959). She carried out two studies in which she developed and tested the Strathclyde Inventory (SI) using data gathered from samples of non-clinical populations.

For the first study she used extracts from Rogers' (1961/1967) writings on the concept of the fully functioning person to develop a pilot 51-item version of the SI spanning six facets of therapeutic change: internal locus of evaluation, openness to experience, self-liking, existential living, acceptance of others and psychological adjustment (Freire, 2007, p.15). The 4-point rating scale tested with the pilot SI ranged from 'fits me pretty much' to 'clearly doesn't fit me'.

Freire recruited 122 trainee and practising counsellors as participants who completed a battery of additional self-report measures chosen to assess convergent and discriminant validity: the Scales for Experiencing Emotions (Behr & Becker, 2002), the Rosenberg Self-Esteem Scale (Crandel, 1973), the Clinical Outcome and Routine Evaluation Outcome Measure (CORE-OM; Evans et al. 2002; Connell et al. 2007), and the Marlowe-Crowne Social Desirability Scale (Strahan & Gerbasi, 1972).

Next, Freire removed items that either intercorrelated with other SI items higher than .7, with CORE-OM items higher than .5 or had loadings of <.5 for a two-factor solution after varimax rotation. Conducting an exploratory factor analysis with the remaining items, she found that a two-factor solution accounted for 43.41% of total variance, indicating a clear separation of items between Factor 1, a positively worded group (23.42% of total variance), and Factor 2, a negatively worded group (19.99% of total variance). She argued that the size and loadings on these two factors suggested that they reflect something distinctive rather than representing an artefact of using positively and negatively worded items (Freire,

2007, p.18). She named Factor 1 *congruence/ experiential fluidity* and Factor 2 *incongruence/ experiential constriction*.

Freire found that the SI had excellent item reliability (Cronbach's alpha = .94; Factor 1 = .92; Factor 2 = .90), and was not substantially associated with social desirability (r = .27). In general, the convergence and discrimination between scores on the SI and the other selected measures were as expected with one significant exception: a high correlation with CORE-OM, indicating that scores on the SI suggested a greater overlap with clinical distress than Freire (2007, p.32) had expected or desired.

Using these results, Freire developed a revised 31-item version and, following feedback from participants, changed the rating scale to a 5-point scale, ranging from 'never' to 'all or most of the time'. Her second study assessed this revised version for temporal consistency and to evaluate convergent and discriminant validity using a different set of selected measures: NEO Five-factor Inventory (NEO-FFI; Costa & McCrae, 1992), Inventory of Interpersonal Problems (IIP – 26-item short form; Maling, Gurtman & Howard, 1995) and the Toronto Alexithymia Scale (TAS; Bagby, Parker & Taylor, 1994). The CORE-OM was also included in order to test if the results of Study 1 would be replicated.

Participants were recruited for two different sets of analyses. Samples 1 and 2 were recruited from two U.S. Midwestern universities: Sample 1 comprised 202 undergraduate freshmen who completed the full set of measures for the convergent validity study; Sample 2 consisted of 67 Masters students in counselling within a broadly humanistic program who took part in the temporal consistency study. Sample 3 was composed of 130 British person-centred counsellors and counselling trainees. They participated in the temporal consistency study and contributed to a subset of the convergent validity study by completing the CORE-OM.

Freire noted significant differences between the SI scores in the three samples. In particular, the mean score for the US freshmen (Sample 1) was

significantly lower than the mean of the other two samples. However, Freire (2007, p.23) concluded that, while the total effect of age and sample appeared to explain about 9% of the total variance, the two variables were confounded and it was impossible to determine from the data which variable was responsible for the difference in the scores. Instead, Freire (2007, p.32) proposed that the variation may be better explained by the possibility that Samples 2 and 3, as humanistic or person-centred trainees and counsellors, had greater self-awareness and/or more familiarity with the language and concepts underpinning person-centred therapy.

Freire confirmed in her second study that there were no remaining items with inter-item correlation larger than .7. Internal consistency remained high (Cronbach's alpha = .94). An exploratory factor analysis once more identified a two-factor solution, this time accounting for 44.67% of the total variance. Again, the separation of items into factors reflected the direction in which the items were worded: Factor 1 (28.3% of total variance) comprised all 19 positively worded items and Factor 2 (16.4% of total variance) was made up of all 12 negatively worded items. Freire reported a Pearson correlation of -.46 suggesting a moderate correlation between the two factors. As a result, she concluded that these two factors replicated the findings of her first study but noted that this was a controversial result because she could not conclude from the data that the two factors were not an artefact of using positively and negatively worded items.

The findings of the convergence validity study were consistent with Freire's theory-based predictions. Freire expanded her set of analyses by conducting Pearson correlations between the two SI factors and the selected measures. For example, there was a significantly higher correlation between the IIP and Factor 2 (incongruence/experiential constriction; r =.48) than between the IIP and Factor 1 (congruence/experiential fluidity; r = -.37). Moreover, while there was a high correlation between the SI and CORE-OM, there were lower correlations between the SI and two sub-domains of the CORE-OM: problems/symptoms (r = -.45) and risk (r = .41). It was proposed that this was consistent with the design of the SI as a "non-pathology oriented measure" (Freire, 2007, p.29). She also noted that the

correlation between SI and CORE-OM was significantly higher for the US freshmen (Sample 1; *r* = -.67) than the British counsellors and trainees (Sample 3; *r* = -.49) and suggested that this finding may indicate that the counsellors and trainees had more ability to differentiate their feelings and greater self-awareness, resulting in them being able to make distinctions in their responses to the two measures (Freire, 2007, p.30).

The temporal consistency of the SI was checked using test-retest data gathered from a subset of Sample 2 (N=45) and Sample 3 (N=32; the counselling trainees). The interval between the two observations ranged between 2 and 10 weeks (mean = 5.22 weeks; mode = 4 weeks). The Pearson correlations between test and retest scores were .79 (the Masters students) and .62 (the counselling trainees). Freire (2007, p.31) argued that the lower test-retest correlation amongst the trainees may reflect the impact of their experiential training, designed to promote self-development and self-awareness, which continued during the interval between tests. She concluded that this could indicate the potential sensitivity of the SI to detecting change experienced through learning activities based on person-centred theory.

Based on her results, Freire (2007) concluded that the SI demonstrated excellent item reliability and adequate temporal consistency when using data collected from UK- and US-based non-clinical populations. She proposed that the high correlation between scores on the SI and CORE-OM may either reflect a limitation of self-report measures in general - that they can be very good at measuring degrees of distress but less able to make conceptual distinctions about the experience of distress - or support the view that using growth-oriented measures make the practice of measuring symptoms of distress redundant.

Further Development of the Strathclyde Inventory

When Freire completed her original work on the SI, she planned to continue the process of validation by conducting a Rasch analysis using the existing data to revise the rating scale and generate a shorter version of the instrument. She proposed that testing the measure with data collected from a clinical population

would follow. However, the opportunity for data collection with a clinical population presented itself before she completed the next stage of her work. When the Strathclyde Counselling & Psychotherapy Research Clinic ('the research clinic') opened its doors in September 2007 it was decided that the 31-item version of the Strathclyde Inventory ('Version 2', developed by Freire for her second study) would be included amongst the outcome measures for both its practice-based (generalist) and social anxiety (specialist) protocols. However, before commencing data collection, Elliott (the Clinic's principal investigator) and Rodgers (the clinic coordinator) revised the wording of some items to improve grammar or because they seemed to contain more than one idea. For example: *I have been able to be spontaneous and genuine* was replaced with *I have been able to be spontaneous*; and *I have felt myself doing things which I could not control at all* was replaced with *I have felt myself doing things that were out of my control* (Freire, 2007; Elliott & Rodgers, 2007). Acknowledging Freire's work in progress, this revised 31-item version was labelled 'Version 4'. Over the next few years, further 'branching' in the development of the Strathclyde Inventory occurred: see Figure 2.1 below.

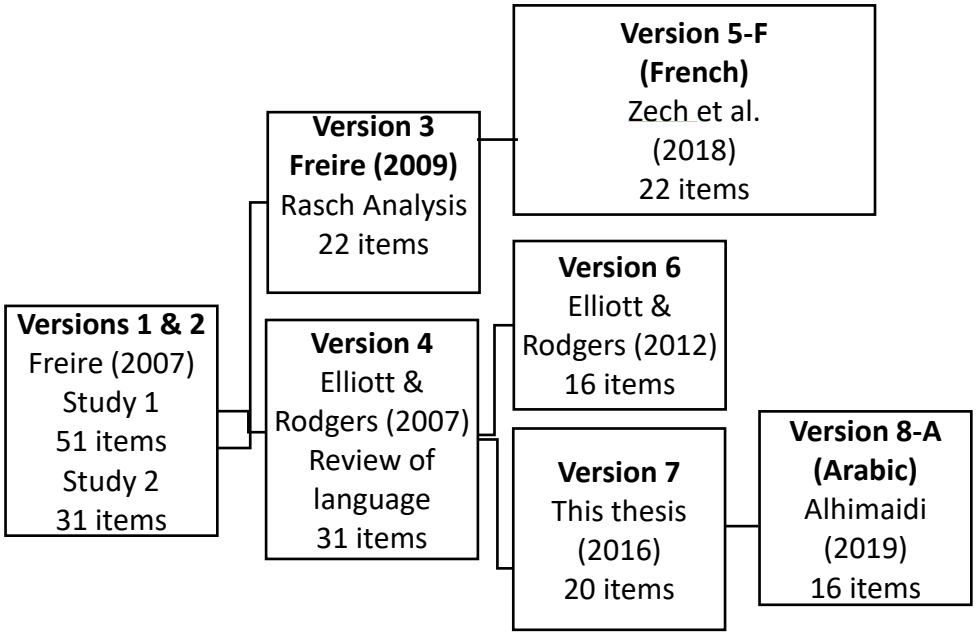


Figure 2.1. Map of the development of the Strathclyde Inventory

Freire concluded her Rasch analysis in 2009, producing a 22-item version (version 3). There is no record of new data being collected on this version. Instead, Freire collaborated with a group of Belgian researchers to produce a French language version of version 3 (labelled 'Version 5-F' in Figure 2.1) that was tested with both clinical and non-clinical populations (Zech et al., 2018). These results are presented in Table 2.1 and discussed below.

Meanwhile, Elliott & Rodgers (2012a) used SI data collected in the research clinic to assess the internal consistency and potential item redundancy of version 4. This was a pragmatic process based on inter-item correlations and exploratory factor analyses, producing a 16-item version ('Version 6'; Elliott & Rodgers, 2012b, 2012c), which was introduced to the research clinic in 2012 as a replacement to version 4. Most recently, a 20-item version of the Strathclyde Inventory was developed at an early stage of this thesis ('Version 7'; see Appendix D), which has been subsequently translated into Arabic, tested with clients attending a university-based counselling service in Saudi Arabia and revised into a 16-item version ('Version 8-A'; Alhimaidi, 2019).

Table 2.1. Internal structure of the SI by version

Version	1	2	4	5-F	6
Sample	NC	NC	C	C & NC	C
N observations	122	399	455	518	455
N items	51	31	31	22	16
Internal consistency (α)	.94	.94	.95	.88	.93
% total variance	43.41	44.67	46.49	42.44	50.79
Factor 1					
N items[1]	19	19	19	14	10
% variance	23.42	28.3	25.53	27.19	29.02
α	.92	.92	.95	.90	.92
Factor 2					
N items[2]	11	12	12	8	6
% variance	19.99	16.4	20.96	15.25	21.78
α	.90	.90	.88	.76	.84

Note. Analyses conducted by: Version 1 (Freire, 2007); Version 2 (Freire, 2007); Version 4 (Elliott & Rodgers, 2012a); Version 5-F (Zech et al., 2018); Version 6 (Elliott & Rodgers, 2012b). NC = non-clinical sample; C = clinical sample. [1] = all items are positively worded; [2] = all items are negatively worded.

Table 2.1 presents an overview of available findings about the internal structure of the Strathclyde Inventory using scores collected from a range of clinical and non-clinical populations across these multiple versions. This overview does not include data for version 7, the development of which will be presented in Chapter 4, nor version 8-A as Alhimaidi (2019) used Rasch analysis (Bond & Fox, 2015) to evaluate the structure of her translated version.

On the whole, Table 2.1 demonstrates very similar findings about the internal consistency and potential factorial structure across these multiple versions of the Strathclyde Inventory. There are two interesting differences. First, the Cronbach's alpha for Version 6 does not decrease as would be expected for a shorter version of an instrument (Barker, Pistrang & Elliott, 2016, p.63), in this case a version with less than one third of items in Version 1 and almost half of the number of items in Versions 2 and 4, indicating that even 16 items may be more than necessary to reliably measure the construct. Second, it is noticeable that the results concerning the total percentage of variance and the percentage of variance explained by Factor 2 for Version 5-F, using data collected from both clinical and non-clinical populations, are closer to those calculated for Versions 1 and 2, both calculated using data collected from non-clinical populations, than Versions 4 and 6, both calculated using data collected from clinical populations. This may be an artefact of the mixed sample or could represent cultural differences.

The design of the study to evaluate the validity of Version 5-F, the French language version of the 22-item Strathclyde Inventory (Zech et al., 2018), was somewhat complex but provided interesting findings. In total, 518 Belgian participants were recruited into the study: four student samples from a large university and four patient samples either from a university hospital (where they were being treated for anxiety, alcohol or cancer) or from a psychiatric clinic. Two of the student samples contributed to a test-retest study, including one sample of undergraduate sophomores randomly assigned to three conditions of counselling training that took place in the period between observations, allowing for comparison of the effect of experiential and non-experiential training methods. The

other two student samples completed the measure on one occasion only. The four patient samples completed the measure twice: on entering their treatment and either at the end of treatment or after a set period of 3 or 6 months (Zech et al., 2018, p.167). The purpose of this data collection was to assess both test-retest consistency and also the measure's sensitivity to change over time in response to a variety of clinical experiences. A range of other instruments designed to measure individual traits, therapeutic attitudes and symptoms were selected by the researchers and completed by various samples according to the design of their particular part in the overall study.

Zech et al. (2018, p.175) claimed that that the data collected using Version 5-F demonstrated that the instrument had very good inter-item consistency and adequate temporal consistency when used with student, trainee and patient populations. Correlations of SI scores with those of other selected instruments indicated convergence and discrimination consistent with their predictions: for example, the SI scores had a high positive correlation with indicators of emotional intelligence and moderate positive correlations with indicators of extraversion and agreeableness and with data collected that indicated the presence of Rogerian therapeutic attitudes during specific encounters with real or simulated clients. They also identified moderate negative correlations with indicators of alexithymia and neuroticism (in students) and symptoms of anxiety and depression (in patients).

They continued to wrestle with the question of whether their results best fit a one- or two-factor model because of the replicated, possibly artefactual, finding of a clear separation of positively and negatively worded items. In the end, based on the moderate correlation (r = .45) they detected between the two factors and their awareness of preliminary findings from the first study reported in this thesis, they proposed "a hybrid model of an over-arching Congruence-Incongruence dimension, with two sub-factors" (Zech et al., 2018, p.176).

Based on their findings, Zech et al. (2018) argued that the measure could distinguish data collected from people who have been diagnosed with a mental health disorder (and theoretically more likely to be incongruent) and data collected

from a non-clinical population: psychology students, psychotherapy trainees and people diagnosed with a physical, not psychological, condition (p.176). In addition, they claimed that the SI was able to detect change among psychology students participating in experiential training and no change in scores for others who participated in no training during the test-retest period or a structured more academic, less experiential counselling skill training. They also found evidence that scores from the SI are sensitive to change resulting from a combination of medical and psychotherapy treatments for anxiety and alcohol patients but not after three months treatment for cancer (p.177).

In reporting limitations of their study and recommendations for future research, the researchers reported one question that they found themselves unable to answer: why did psychiatric inpatients register unexpectedly high scores on the SI (indicating greater congruence) that remained stable when measured six months later having received treatment as usual in the interim? This presents an interesting challenge to researchers investigating the validity of SI scores: is it possible that incongruence might lead to someone selecting responses on the SI with higher ratings due to a lack of congruence? This potential finding would fit with the theory of congruence/incongruence but may be harmful for the endeavour to develop a measure of the construct.

Measuring Change using the Strathclyde Inventory with a Clinical Population

Meanwhile, data collection to test the ability of the English language SI to be used as a measure of change in a clinical setting was ongoing at the research clinic. The process began in Autumn 2007 when Version 4 of the SI was included in the battery of outcome measures used in the two protocols developed within the research clinic: the general 'Practice-Based' (PB) protocol and the specialist 'Social Anxiety' (SA) protocol.

Since then, two studies have been carried out to analyse the earliest outcome data gathered in the clinic: Tashiro (2011), who analysed the outcome results of the first 43 participants in the PB protocol, and Elliott, Rodgers and Stephen (2013), who analysed the outcome results of the first 53 participants in the

SA protocol. In both studies, a large pre-post effect size was found on the SI. In Table 2.3 the results of these studies are presented along with the equivalent results found in each clinical sample contained within the Zech et al. (2018) study and those of the participants in the Alhimaidi (2019) study.

Table 2.2. Pre-post scores and effect sizes for studies using SI to measure change with a clinical population

	Context for data collection	Pre-Therapy			Post-Therapy			Effect Size (*d*)
		n	m	sd	n	m	sd	
Tashiro (2011)	Practice-Based protocol	39	1.90	.64	39	2.73	.82	1.13
Elliott et al (2013)	Social Anxiety protocol	50	1.75	.51	42	2.39	.53	1.25
Zech et al. (2018)	Anxiety clinic	9	2.22	.53	9	2.37	.28	.35
	Alcohol withdrawal	56	2.34	.72	56	2.93	.64	.87
	Psychiatric ward	10	2.71	.42	10	2.77	.59	.12
Alhimaidi (2019)	University counselling service	38	2.22	.79	34	2.77	.54	.79

Notes. n = number, m = mean, sd = standard deviation; *d* = Cohen's *d*.

As Table 2.2 shows, there is a small but growing number of studies in which data has been collected on the SI from a range of clinical populations to measure change resulting from a variety of interventions. It is interesting to note that the largest effect sizes so far have been calculated on the scores collected from clients participating in person-centred therapy (Tashiro, 2011) and either person-centred therapy or emotion-focused therapy (Elliott et al, 2013) at the research clinic. It is possible that this indicates that greater change in scores on the SI (aggregated across clients) may be experienced as a result of participating in person-centred and emotion-focused therapies. However it must also be acknowledged that, first, the pre-scores for both of these samples were the lowest of all of the studies being

compared and therefore there was greater scope for increased scores, and, second, the duration of treatment may well have been longer, especially for the participants in Tashiro's study as the research clinic's PB protocol offered up to 40 sessions of person-centred therapy. In contrast, participants in the three week long alcohol withdrawal program, reported by Zech et al. (2018), accessed one week of medical and psychological treatment, one week at home, then one final week of treatment, while participants in the Alhimaidi (2019) study accessed short-term counselling (4-6 sessions) offered by counsellors trained in a variety of therapies including cognitive behavioural therapy and person-centred therapy. It must also be noted that these differences in pre-therapy scores may be linked to the use of different language versions or indeed reflect a different understanding or experience of the construct being measured amongst Scottish, Belgian and Saudi Arabian participants.

In addition to investigating change in scores on the SI, Tashiro (2011) carried out a detailed study of the relationship between the scores collected from the participants on all three outcome measures used in the PB protocol – the SI, the CORE-OM and the Personal Questionnaire (PQ; Elliott et al., 2016). Consistent with Freire (2007), she found a strong correlation between ratings on the SI and CORE-OM, a measure of general distress, but a less strong association between scores on the SI and the PQ, an individualised measure of specific personal difficulties identified by the client. She noted that the pre-therapy ratings on the SI were more varied or scattered than those of the other measures, and that these were more inclined to begin in the non-clinical range at pre-therapy than scores on the other two instruments: SI = 20.5%; CORE-OM = 12.5%; PQ = 4.7%.

Tashiro (2011) proceeded to examine the clinical significance and reliability of change (Jacobson & Truax, 1991) reported by each participant and discovered that 51.6% of the sample who had been in the clinical range on the SI at pre-therapy reported clinically significant reliable change. 43.6% of all clients reported non-reliable change. This group included 38.7% of those participants whose scores had been in the clinical range at pre-therapy and 62.5% of those whose scores had been non-clinical at pre-therapy. No reliable deterioration was identified in any

participant's scores. Tashiro proposed two hypotheses that may explain the somewhat larger proportion of non-reliable change reported on the SI compared with the other two outcome measures: (a) that items in the SI reflected aspects of experience where change did not easily occur during the treatment period compared to the areas targeted by the other two instruments; and (b) the size of pre-therapy clinical case ratio affected the size of all-case improver ratio, suggesting that the measure may not be of initial relevance to clients experiencing poor psychological health.

Tashiro's findings raise interesting questions about how clients may perceive themselves in relation to the items contained in the SI, especially at the beginning of the therapeutic process. In addition, if a client is relatively incongruent at pre-therapy, therefore low in self-awareness, then it is possible that this will impact on the self-report data collected by the SI at pre-therapy. It may be that as the client proceeds in therapy, and self-awareness increases, they may become more aware of their lack of congruence and this may register in their SI scores completed at mid-therapy points. Therefore, the pattern of change in scores through the process of therapy needs to be investigated.

However, it is also important to gain some perspective on these results by comparing Tashiro's (2011) findings about the SI with those she made about the scores from the same sample of clients collected using the other two instruments, CORE-OM and PQ. Scores on the SI suggested that 51.6% of clients experienced clinically significant change compared with 54.3% of clients, according to their CORE-OM scores, and 56.1% of clients, based on their PQ scores. Alongside this, while Tashiro (2011) found that 38.7% of clients whose SI scores indicated that they started in the clinical range and experienced non-reliable change, this figure for clients according to their CORE-OM scores equated to 31.4% and for PQ scores was 34.1%. In both examples, this comparison suggests similar clinical outcomes across measures on these indicators. Furthermore, Tashiro (2011, p.33) noted comparability in the ratios of clinically significant change when benchmarking her results with two large-scale studies using outcome data collected in naturalistic

settings (Stiles et al., 2006; Stiles, Barkham, Mellor-Clark & Connell, 2008). Tashiro (2011, p.65) proposed a number of administration and design issues that might influence the differences in scores and their interpretations between the three instruments including: (1) the PQ may have an advantage in being created by the participant and completed more frequently; (2) the SI is usually completed after the CORE-OM; (3) the SI asks participants to reflect over the past month, whereas the PQ and CORE-OM ask participants to reflect back over the past week only; and (4) the PQ has a 7-point scale which may allow more accurate reporting compared with the 5-point scales used by the CORE-OM and SI.

Calculation of Clinical Significance on the Strathclyde Inventory

Jacobson and Truax (1991) introduced statistical methods to calculate clinical significance in psychotherapy outcome measurement. They argued that statistical significance and effect size in aggregated outcome data may bear little resemblance to the change experienced by the individual. In contrast, the concept of *clinical significance* captured the idea of a person entering therapy as a member of a 'dysfunctional' population and leaving therapy as part of a 'functional' population. They proposed that a cut-off score could be calculated that identified the threshold between the two populations (p.13). In addition, they noted that the change from dysfunctional to functional required another test in order to be credible and recommended the calculation of a *reliable change index (RCI)* that would establish how much change in score would be required to ensure that it was 'real' change, not measurement error (p.14). As a result, an important function in the development of measures for use in counselling is the calculation of these metrics for the instrument. Jacobson and Truax proposed three alternative methods for calculating a cut-off score that depended on whether data on the instrument was available from a 'dysfunctional' population (criterion A), a 'functional' population (criterion B), or both (criterion C).

Elliott (personal communication, 2019) has offered a set of calculations of a cut-off for clinically significant change and a reliable change index for use by researchers using the Strathclyde Inventory (see Table 2.3).

Table 2.3. Clinical Significance Cut-off Score and RCI using data from previous studies.

SI Version	Study	Cut-off	RCI[a]	RCI[b]	Criterion
Version 2	Freire (2007)	1.69	.70	.50	B
Version 4	Folkes-Skinner (2011)	2.45	.60	.40	C
Version 5	Zech et al. (2018)	2.52	1.02	.67	C

Notes. [a] = $p < .05$; [b] = $p < .2$

The cut-off score calculated using Jacobson and Truax's (1991) criterion B with data collected from a non-clinical population by Freire (2007) appears to be out of step with the two calculations in which Jacobson & Truax's criterion C were used where both clinical and non-clinical data was available. The two RCIs calculated by Zech et al. (2018) are noticeably higher than the others. This difference is caused by the lower test-retest score obtained in that study: .59 compared to scores of .79 in the other two studies.

Jacobson and Truax (1991, p.14) recommended that standardised clinical significance and RCI scores based on normative data are calculated by aggregating scores across studies. Standardised scores for the SI have not yet been established.

Chapter Summary

In this chapter I have presented an overview of self-report instruments that have been developed to measure constructs associated with the type of psychological growth, congruent functioning, envisaged by person-centred theory and therefore have the potential to be used as theoretically-congruent outcome measures in PCT. In the second half, I described the development of the Strathclyde Inventory (SI), the instrument that provides the basis for this thesis. Previous studies of data collected on the SI and its French language version (SI-5F; Zech et al, 2018) have provided credible preliminary evidence of its validity based on data collected from non-clinical and clinical populations, while also raising several questions relating to its internal structure, its ability to distinguish congruent functioning from symptomatic experiences of distress, and its usefulness as an outcome measure in therapy.

Based on this review of existing evidence, I anticipated that my own study, using data collected from a large sample of the research clinic's participants, would replicate previous results and find that the SI was able to demonstrate a high degree of internal consistency, reliability/precision, and sensitivity to change over the course of therapy. However, I planned to go further by using a range of methods to investigate the evidence for accepting the SI as a valid measure of congruent functioning and, by extension, using it as an outcome measure in PCT, and in so doing develop a more precise definition of congruent functioning.

Chapter 3: Measurement in Counselling

Chapter Overview

In this chapter, I present the epistemological, methodological and ethical landscape in which research using measurement within the field of counselling takes place. I begin with an introduction to contemporary understanding of measurement in the psychological sciences and contrast that with the more common view of measurement held by counsellors, shaped by the epistemology that underpins person-centred and other humanistic approaches to counselling. Next, I outline the influence of measurement within counselling today and highlight the challenges to this paradigm, in particular methodological and ethical concerns. I summarise the first section of this chapter by making a case for the informed use of measurement in counselling as an essential part of *methodological pluralism* (e.g. Barker et al., 2016) in which the assumptions, strengths and limitations of measurement are recognised. In the final section of this chapter I describe the approach to measurement that I have adopted in this study.

Why is Measurement Important?

There has always been a drive to quantify and measure the physical world, whether for trading, mapping the landscape, or assessing the environment, and therefore simultaneously a need to develop methods for doing so.

Early psychologists brought the same curiosity to their studies and for many, even in those early days, there was no question that psychological attributes could be measured. Slaney (2017, pp.30-31) described the impetus that the publication of Darwin's *Origin of the Species* in 1859 gave to the idea that individual variation of traits and characteristics within a species could be identified and compared, and the development of numerical methods to analyse the data collected. This activity was driven by an ontological belief that such attributes existed in a way that could be quantified and measured: a scientific realism in which there exists a 'real' world which consists of 'things' that are independent of humans' ability to perceive or reason about them; a philosophical positivism that the purpose of science is to

develop progressively more accurate knowledge of the world based on observation; and that the truth of scientific theories depends on the degree of correspondence between them and the real world they seek to describe (Slaney, 2017, p.153). This outlook was operationalised in the true score model underpinning classical test theory, in which it was conceptualised that an individual's observed test score is made up of a true score and error (Slaney, 2017, p.32). As a result, the main focus in measurement at this time was the identification and elimination of measurement error through increasing the reliability of measures.

Since then, there has been considerable debate about the nature of psychological attributes, including how (and if) they can be known and measured. As a result, modern test theory is based on the perspective that theoretical constructs can be identified and defined to represent unobservable psychological attributes. According to item response theory, differences between individuals in relation to these latent traits is reflected in their responses to test items, providing indirect empirical data that can be analysed using mathematical models.

The contemporary view of measurement in the psychological sciences is captured in the *Standards for Educational and Psychological Testing* ("the Standards"; originally, the *Technical Recommendations for Psychological Tests and Diagnostic Technique*). The purpose of the Standards is "to promote the sound and ethical use of tests and to provide a basis for evaluating the quality of testing practices" (Plake & Wise, 2014, p.4) and is aimed at a broad audience: professionals and graduate students in the field of educational and psychological measurement, policy makers, and test users, not only within the USA but also an international audience (Plake & Wise, 2014, p.6). According to the Standards, the three foundations of testing are *validity*, *reliability/ precision and errors of measurement*, and *fairness*.

Validity

The latest version of the Standards defines validity as "the degree to which evidence and theory support the interpretations of test scores for proposed uses of tests" (AERA, APA & NCME, 2014, p.11). Since Messick (1995), the Standards have

advocated a unified concept of validity, proposing that criterion-oriented validity (i.e. predictive validity and concurrent validity), content validity, and construct validity are aspects of the same process. Consolidating this position, the 2014 edition of the Standards recommends the comprehensive investigation of five sources of validity evidence: test content (Standard 1.11; p.26), response processes (Standard 1.12), internal structure (Standards 1.13 – 1.15; pp.26-7), related variables (Standards 1.17 – 1.19; p.28), and intended and unintended consequences (Standards 1.20 - 1.25; pp.29-30).

Adopting Messick's (1995) perspective that validity is not a property of the test itself but of the *meaning* of the test scores, the Standards require that test developers should provide a rationale for, and evidence that supports, the intended interpretation of test scores for each way in which the instrument could be used and with each population (Standards 1.0 & 1.1; p.23), underlining that producing evidence of validity is an ongoing process as the potential uses of the instrument with an expanding range of populations continue to develop.

Reliability/Precision and Errors of Measurement

The latest edition of the Standards introduced the compound term reliability/precision, defined as "the more general notion of consistency of the scores across instances of the testing procedure" (p.33) in order to distinguish this broad idea from the more specific concept of reliability coefficients that has traditionally been applied in classic test theory. This broadening focus implies that it is the *dependability* of the scores that provides the evidence for decisions made as a consequence of the testing. The Standards emphasise that the greater the impact on the life of the test-taker that the decision made on the basis of testing is likely to have, the greater the responsibility on test developers to ensure the reliability/precision of the test. "More modest" reliability/precision is acceptable for tests leading to decisions that are less consequential or are made within the context of a wide range of evidence that includes test scores (AERA et al., 2014, p.33). As Elliot (2015, p.680) commented "reliability is no longer solely a precondition to validity; rather, reliability/ precision is evidence of fairness".

The Standards distinguish between random errors of measurement and systematic errors. Random errors are seen as unpredictable fluctuations that could have their source in the test-takers (e.g. motivation, interest, attention, inconsistency in application) or in the test conditions. In contrast, systematic errors are understood to affect test-takers' scores in a consistent manner and are more likely to affect the validity of the scores rather than reliability/precision, on the grounds of fairness. (AERA et al., 2014, p.36).

Fairness in Testing

In the 2014 Standards, fairness was given equal weighting with validity and reliability/precision and errors of measurement, as one of the foundations of sound and ethical testing practice. Indeed, Elliot (2015, p.678) remarked "If the five previous editions of the Standards have been about validity, then the present edition is about fairness".

The Standards describe fairness in testing as "a desirable social goal", noting that there is "no single technical meaning and [the term] is used in many different ways in public discourse" (AERA et al., 2014, p.49). Therefore, the Standards propose their own definition:

> A test that is fair within the meaning of the *Standards* reflects the same construct(s) for all test takers, and scores from it have the same meaning for all individuals in the intended population; a fair test does not advantage or disadvantage some individuals because of characteristics irrelevant to the intended construct (AERA et al., 2014, p.50).

Lack of fairness (e.g. measurement bias, inaccessibility) is understood to be a threat to the validity of test score interpretation. Potential barriers to fairness in testing are likely to differ from subgroup to subgroup and between individuals within subgroups.

The Standards use the term *predictive bias* to describe evidence that "differences exist in the patterns of association between test scores and other variables for different groups, bringing with it concerns about bias in the inferences

drawn from the use of test scores" (AERA et al., 2014, p.51). Differences between subgroups can be investigated through a potential range of analyses including differential item functioning, differential test functioning and regression analysis.

An important development in the 2014 Standards is the broadening of the principle of fairness so that it applies to *all* test-takers, not only those with disabilities or diverse linguistic backgrounds as had been the case in the previous edition. It is understood that this acknowledgment of the general need to focus on fairness in testing has arisen because of the increasingly important decisions about the lives of individuals that are being made on the basis of test scores (Plake & Wise, 2014, p.10). This is particularly so in the field of education but is also relevant in psychological testing and specifically in counselling, for example, when test scores are used to determine whether an individual can access treatment (e.g. Evans, 2019).

Plake & Wise (2014) have stated that the acceptability of a test should not rest on the "literal satisfaction of every standard… [or] be determined by using a checklist" (p.11). They argue that evaluating acceptability involves professional judgment and the degree to which the *intent* of the standards has been satisfied, amongst other issues. It is clear, therefore, that the Standards, while not a checklist, present an evaluative framework to lead and support a philosophy and practice of contemporary measure development that is as equally relevant and important for measurement in counselling as in other disciplines.

Is Measurement Important in Counselling?

Traditionally, the importance of research in counselling has been rated more highly by researchers than practitioners. In 1986, Morrow-Bradley & Elliott conducted a survey of American psychotherapists, mainly working in private practice, to investigate their use of research and their perception of problems with existing psychotherapy research. They discovered that most participants rated as 'minimal' or 'some' the degree to which research findings influenced the way that they did psychotherapy. Indeed, participants reported that the most important

source of information that they found useful for practice was their own ongoing experience with clients. Therapists who used research less were found to be more critical of it. The main criticisms reported by participants were that research tended to treat therapists as interchangeable, that studies were rarely designed to try to incorporate the complexities of psychotherapy, and that practical, relevant, and scientifically sound measures of psychological change were often unavailable.

Recent research suggests that similar attitudes and practices continue amongst practitioners today. In 2014, Gyani, Shafron, Miles & Rose reported the results of an online survey of 736 therapists in the UK and found that the most used source for routine clinical decision-making was supervision, followed by personal experience with clients, clinical guidelines, peer discussion, clinical case observations, outcome measures, case studies, controlled trials, followed by popular books (p.206). The researchers found that theoretical orientation and work setting were influencing factors: cognitive behavioural therapists and those participants who worked in the National Health Service (NHS) had more favourable attitudes toward research and used research more frequently when making clinical decisions than other therapists. They also discovered, replicating the findings of Hatfield and Ogles (2007), that cognitive behavioural practitioners used outcome measures more often than other therapists.

Therefore, it appears that attitudes towards research amongst counselling practitioners may depend on theoretical orientation, the context in which they work, and, perhaps, their own involvement in data collection through the use of outcome measures within their practice. However, many, including Bondi and Fewell (2016), have contended that the difference goes much deeper than this. For example, Bondi and Fewell argued that the nature of the research that has acquired prominence in the field of counselling and psychotherapy "embodies the very antithesis of what we value most highly in our approach to training and practice" (p.4). Drawing on broader social science perspectives concerned with human meaning and feminist critiques of positivism, they highlighted the difference between research conducted from *experience-near* (i.e. close to the feelings and

experiences of the people taking part; akin to reflective practice) and *experience-far* (i.e. at a distance from the feelings and experiences of those being studied; technically-driven, authoritative knowledge) positions (p.6) and set out three interconnected reasons that explained why the positivism of the natural sciences cannot translate to the human sciences: first, the multiple layers of interpretation inherent in the reflexive process of people studying people, captured by the concept of the *double hermeneutic* (Giddens, 1984); second, the hopelessness of maintaining objective detachment as a human, with "insider status" (p.27), studying the human experience; and, third, the impossibility of separating the focus of study from its context without losing essential aspects of their experience (pp.25-30).

There are many well-articulated criticisms within the person-centred literature, both thoughtful and acerbic (e.g. Hilton & Prior, 2018; Lee, 2018), of research conducted from an experience-far perspective. It is no wonder, therefore, that many counsellors struggle to see relevance for their practice in studies that might be described as experience-far approaches to research, including measurement.

The Influence of Measurement in Counselling

Nevertheless, while the challenge to persuade practitioners of the potential contribution of research and in particular measurement continues, the influence of measurement in the field of counselling today is growing. Next, I present the three main areas in which measurement has impact: *theory development*, *evidence-based practice*, and *routine outcome monitoring.*

Theory Development

Using measurement as a means for theory development in counselling is more commonly the domain of the researcher. This is not to say that only researchers develop counselling theory but that such development does not usually involve the use of measures and has a different epistemological basis, as described above. Contemporary versions of Rogers' original counselling research clinic can be found at various universities in the UK and beyond.

The philosophical foundations for using measurement to extend theory can be found in Cronbach & Meehl's (1955, p.290) description of the principles of a *nomological net* or *network*:

> When a construct is fairly new, there may be few specifiable associations by which to pin down the concept. As research proceeds, the construct sends out roots in many directions, which attach it to more and more facts or other constructs (Cronbach & Meehl, 1955, p.291).

Cronbach and Meehl (1955, p.290) explained that a nomological net is an "interlocking system of laws" in which it can be demonstrated that observable properties relate to each other, theoretical constructs relate to observations, and different theoretical constructs relate to each other. They argued that for a construct to be considered "scientifically admissible", at least some of its nomological net must be based on the results of observation, which either extends the network or increases the strength of the evidence underpinning inferences of links within the chain.

The idea of the nomologicial net emphasises the long-term commitment involved in theory development, involving an iterative cycle of induction, deduction and abduction, described by Rennie (2000) as *methodical hermeneutics*. The process requires the identification of a theoretically-relevant phenomenon that can be defined as a construct, and a rigorous programme, first, to develop an instrument that had credibility as a measure of that construct, second, to collect and interpret the meaning of the data collected on the instrument, and then, to replicate or identify and test new ideas about the construct. Also essential are knowledge, expertise, time, the ability to collect a significant amount of data and the adoption of a philosophical stance that can negotiate the tension between the epistemology underpinning measurement and that of most humanistic counsellors. This last attribute is essential: if the researcher cannot investigate and interpret their data and communicate their findings in a way that makes sense and is relevant to practitioners then they will reinforce, rather than bridge, the gap between research and practice.

An example of a long-standing and ongoing programme of measurement-related theory development that has influenced decisions in practice, and is relevant to this study, is research into the amount of therapy necessary for the client to gain a positive outcome. First, Howard, Kopta, Krause, & Orlinsky, (1986) identified a *dose-effect* relationship between the number of sessions that participants accessed in therapy and ratings of improvement across therapy. Their results showed that, irrespective of their actual duration in therapy, 48% to 58% of participants would be expected to have improved after eight sessions, about 75% after 26 sessions (six months), and about 85% after 52 sessions (one year). They also compared rates of change between participants according to types of diagnoses and noted differences between them. The implications of this study were significant: not only providing evidence of 'treatment response' in therapy, going beyond a simple examination of effectiveness based on pre-post data, but also in providing preliminary guiding information for practitioners about how much therapy to offer.

Further refinement of the dose-effect model occurred over the next twenty years, creating what Barkham et al. (2006) described as "a lively and substantial research literature" (p.160). Challenging a fundamental assumption underpinning the dose-effect model, Barkham et al. proposed a different interpretation for the negatively accelerating curve that had been understood to represent a diminishing return on sessions for individual clients beyond a certain level of improvement. Instead, according to Barkham et al., this curve could reflect the phenomenon of quickly improving clients leaving therapy, having reached a *good enough level of improvement* (GEL), resulting in estimates for improvement at later sessions being based on data from gradually reducing groups of more slowly responding clients. Barkham et al. suggested that this GEL model indicated that the relationship between duration of therapy and degree of improvement was "mutually regulated" (p.161) by clients ending therapy when their sense of their improvement felt satisfactory. The GEL model was supported by evidence that 71.7% of the 1442 participants in their study who began therapy above the clinical cut-off for the

chosen outcome measure had a planned ending and demonstrated reliable and clinically significant change (Jacobson & Truax, 1991) according to their post-therapy scores. The development of the GEL model promoted the idea that achieving an acceptable outcome was a subjective assessment, likely to require individuals to access different lengths of time in therapy, thereby creating less certainty for service providers wishing to control their limited resources by offering a standard number of therapy sessions.

Both models assumed a steady rate of improvement across therapy, albeit varying according to individual factors. The *shape of change* over the course of therapy became an area of interest for researchers, in particular the phenomenon of *early change*. For example, Stulz et al. (2007) analysed self-report data collected from 192 clients on at least three occasions during the first six sessions of NHS counselling. They identified five distinct client groupings that they named according to characteristics of the shape of early change indicated by their scores: high initial impairment, low initial impairment, early improvement, medium impairment with continuous treatment progress, and medium impairment with discontinuous treatment progress. These shapes of early change were found to be associated with different outcomes and durations of therapy, some surprising: for example, a higher proportion of clients in the *discontinuous progress* group ended therapy with good outcomes, according to their scores, compared to the *continuous progress* group. As Stulz et al. noted, this finding offered evidence that worsening scores in the early stages of therapy did not necessarily result in a poor outcome for the client (p.871).

This research strand continued to develop with specific focus on the relationship between early change and outcome for participants with specific diagnoses: for example, depression (Lutz, Stulz & Köck, 2009) and panic (Lutz et al., 2014). Cumulative results demonstrated that early change patterns are more accurate at predicting outcome than the type of therapy provided, strengthening the growing field of research into *feedback monitoring systems* (e.g. Lambert et al., 2002) that investigate the impact of providing therapists with feedback on client

progress during the therapeutic process in order to address potential issues indicated by lack of change in weekly scores.

Nevertheless, the work of Owen et al. (2015) expanded the proposition that good outcome is not necessarily determined by early change. Using data collected from 10,854 participants, Owen et al. examined trajectories across longer periods of therapy (5 to 25 sessions) and identified at least three distinct patterns of change across therapy: *early and late*, *worse before better*, and *slow and steady*. The most unexpected finding was the second period of improvement detected for participants who had experienced early change, indicating the potential benefit of continuing in therapy beyond the initial experience of rapid change. Overall, Owen et al.'s finding emphasised that a variety of pathways can lead to positive change over the course of therapy and that clients whose progress is more gradual are likely to benefit from more sessions.

As these examples show, measurement has made a fundamental contribution to this long-standing programme. Though this work, our understanding of the relationship between one aspect of the context required (i.e. access to sufficient sessions) and outcome in therapy has developed, leading to recognition of a variety of patterns of change that enable clients to use and benefit from their experience in counselling.

Evidence-based Practice

The paradigm of *evidence-based practice*, as a system for producing methodologically rigorous and robust research to inform policy and practice in medicine, was adopted into the field of psychological therapy in the 1980s (Barkham & Mellor-Clark, 2003). The focus is *efficacy*, a replication of the medical model approach: in other words, identifying if a specific intervention has the intended effect, and, if so, how much of the intervention is required, and what side effects may result. The vehicle for the production of evidence-based practice is the randomised control trial (RCT), in which the emphasis on randomisation, manualisation, strict inclusion/exclusion criteria, and the use of a control group for comparison is designed to strengthen the internal validity of the results.

Two products of person-centred and experiential research that have been developed to enable PCT to participate within the paradigm of evidence-based practice are the IAPT-approved version of PCT, Counselling for Depression (CfD; Proctor & Hayes, 2017), and the adherence and competence measure, Person-Centred and Experiential Process Scale (PCEPS: Freire, Elliott & Westwell, 2014).

Critics of evidence-based practice, and the RCT in particular, have argued that this attempt to adopt a sterile laboratory-style metaphor to conduct counselling research significantly weakens the external validity (i.e. generalisability) of the results, rendering the results of evidence-based practice largely irrelevant for practitioners (e.g. Evans, Connell, Barkham, Marshall & Mellor-Clark, 2003; Kennedy-Martin, Curtis, Faries, Robinson & Johnston, 2015). An innovation to address some of the issues identified with the use of the RCT as the basis for evidence-based practice in counselling and psychotherapy has been the introduction of the *practical RCT* (e.g. Freire et al.,2015; Farr, Di Malta & Cooper, 2019), an attempt to conduct a small-scale research programme that follows an RCT design, within routine practice. However, this is far from easy as Farr et al. (2019) outlined in their description of the recent failure of a practical RCT, citing real world issues such as staff turnover, additional demands on (in this case, volunteer) practitioners, poor communication about research procedures, and lack of motivation and engagement by practitioners.

Practice-based Evidence

The paradigm of *practice-based evidence*, as a way of providing evidence derived from routine practice (Barkham & Mellor-Clark, 2003), was an early response to criticisms about the relevance of evidence-based practice for therapists in the real world. In contrast to evidence-based practice, the focus is *effectiveness,* maintaining an acceptable level of quality when working with the diverse range of clients coming to therapy. Margison et al. (2000) outlined several ways in which measurement can be used to bridge the research-practice or "efficacy-effectiveness" (p.129) gap, most prominently, *routine outcome monitoring* (ROM). Mellor-Clark, Cross, Macdonald & Skjulsvik (2016) defined ROM in psychological

therapy as “the practice of inviting users of psychological services to complete standardized questionnaires to help profile progress (or otherwise) in treatment” (pp.279-280). These self-report questionnaires are considered to be a quick and cheap method for collecting data that can be analysed and used for a variety of purposes.

In contrast to RCT research, ROM privileges external validity, as it reflects routine practice, but at the expense of internal validity, as it is carried out in a context that makes it much more difficult to distinguish between the many variables that may have contributed to an individual’s outcome (represented by their post-therapy score) at the end of therapy. Indeed, large samples of data are required to produce the statistical power necessary to approach the robustness of RCT design (Barkham & Mellor-Clark, 2003). As a result, advocates for ROM support the development of *practitioner research networks* (e.g. Barkham & Mellor-Clark, 2003). One example, the CORE national practice-based data set (Evans et al., 2003), has been the basis for studies of the effectiveness of counselling in naturalistic settings based on large-scale ROM data (e.g. Stiles et al., 2006; Stiles et al., 2008).

What does ROM measure? As Evans et al. (2003) reported, the purpose of ROM is not only collecting data but also enabling it be used in a meaningful way, and therefore requires the identification of an appropriate and relevant measurement tool and the development of systems for reporting the results, not only within the organisation but to the wider field. Originally ROM required pre- and post-therapy data collection only - at the first and final counselling session - however in recent years the collection of ROM data at every session has been promoted as best practice (Mellor-Clark et al., 2016). It is argued that this increases the chance of collecting final session data in the event that the end of counselling is unplanned, although Mellor-Clark et al. (2016) noted that: “irrespective of the frequency of measurement, inefficient measurement administration processes coupled with high rates of early attrition from treatment [has] consistently compromised the volume (and hence reliability) of outcomes data” (p.282).

How are the products of ROM used? Evans et al. (2003) reported that the development of the CORE national practice-based dataset was to provide a benchmark for practitioners and services: "not standards set by government or other organizations [nor] generated in response to political or managerial target setting. They simply reflect practice as it is in the range of routine clinical services participating in the data collection" (p.385). At the same time, they noted the need for counsellors and service managers to be able to interpret their results in the context of local conditions and argued against the development of "ranking or 'star chart' systems" (p.386).

Lucock et al. (2003) described a practice-based service model in which therapists received individual and grouped data and graphs as feedback on their clients' outcomes, in which they were beginning to develop the practice of discussing the findings with the therapists as a way of developing an understanding of the data from a clinical perspective, promoting *evidence-based reflective practice* (p.392) and fostering increased quality of service. They argued that the main challenge is creating an environment in which ROM is experienced by staff in a non-threatening way, noting that this is essential for the process to be maintained on an ongoing basis and result in service improvements.

Boswell, Kraus, Miller and Lambert (2015) highlighted the benefits of using ROM as a means of tracking client progress during therapy by increasing therapists' awareness and therefore ability to respond to clients whose scores predict negative outcome. Research into the effect of using ROM-based feedback systems to track progress on therapy outcome has been conducted, for example in the recently reported study by Brattland et al. (2018) in which an RCT design was used to investigate the impact of tracking progress using ROM on outcomes in a mental health clinic over a four year period. The results showed that clients in the ROM condition were 2.5 times more likely to demonstrate reliable improvement than clients receiving treatment as usual, even when therapist variability and client pre-therapy distress levels were controlled. The effect of tracking progress via ROM on

outcome increased month on month across the implementation period, suggesting that therapists became more effective in its use over time.

However, researchers have highlighted difficulties in implementing ROM. For example, Boswell et al (2015) identified barriers caused by practical issues, such as financial and time burdens, multiple stakeholders (e.g. clients, therapists, administrators, funders) with different needs, and staff turnover, and also resulting from philosophical concerns that included practitioner scepticism about the validity of outcome assessment, and fear and distrust about how the data will be used. They recommended that direct experience of the benefits of using ROM, and the sharing of those experiences among therapists is the most effective way to reduce practitioners' concerns. They also promoted the use of simple and minimally disruptive data collection procedures that offer some degree of flexibility for practitioners as well as a collaborative and transparent approach in which ROM is not imposed but rather co-constructed with practitioners.

Challenges to Measurement in Counselling

I have highlighted several potential challenges to measurement in counselling in the previous section. These challenges are based on methodological and ethical concerns.

Can We Measure 'Change' in Counselling?

Although the idea of measuring change over the course of therapy appears to be a simple and straightforward process achieved by calculating the difference between two scores, this is not the case. Cronbach & Furby (1970) argued that measuring change by calculating the raw difference between pre- and post-test scores was "rarely useful" (p.68). Two reasons that they provided are particularly relevant in relation to measurement in counselling. First, that pre- and post-therapy scores are 'linked' (i.e. non-independent) because the same person has provided the data, albeit on different occasions. Instead, they recommended the calculation of a residualised score as it controlled for the pre-test score and therefore identified test-takers whose scores have changed more or less than

expected, providing evidence for an inference about outcome. Second, Cronbach and Furby emphasised that as "change is multivariate in nature" (p.76), scores from the same person at different times may not represent the same psychological process. In other words, a client may be working on one aspect of their experience early in therapy, and quite a different aspect later on. Therefore, any change in scores that the client has recorded may not be representative of the construct being measured, and conversely any raw difference in scores may not necessarily reflect the client's experience in therapy.

More recently, Langkaas, Wampold and Hoffart (2018) recommended a change to the reference model underpinning the interpretation of progress within ROM. They noted five different ways to model difference in scores (i.e. observed difference, detected difference, predicted difference, attainment difference, and induced difference). They proposed that *induced difference* (the difference between observed progress and that which would have occurred naturally without intervention) is a more appropriate model for measuring effectiveness in clinical practice. In addition, they argued that the model underpinning current ROM systems emphasised positive change as an indication of progress and is unable to distinguish cases in which no change might be a therapeutic goal or, indeed, when getting worse initially may still lead to a positive outcome, citing the findings of Owen et al (2015).

These two perspectives, almost fifty years apart, highlight the complexities inherent in the kind of change involved in counselling and the limitations of using measurement as a means of representing it, especially in ROM when the responsibility for calculating and interpreting the meaning of change in scores is typically in the hands of non-statisticians, whether counsellors or administrators, who simply apply an observed (raw) difference model.

Issues with self-report measures. In addition to the limitations associated with the models used to calculate difference in scores, there are recognised limitations inherent in the use of self-report instruments.

As McLeod (2001) described, self-report instruments were originally developed in order to measure attributes that were understood to be stable, and designed to be used as one-off assessment tools that could be administered in large groups to support screening or selection decisions. Therefore, the question of whether using a self-report instrument to measure change was an appropriate or sound method was not considered. Reviewing the subsequent literature, McLeod noted the well-known impact of social desirability on item response and identified this as problematic when using self-report instruments as a measure of therapeutic change because: first, social cues are implicit in the format and language of instruments and, second, the practice of completing them on regular occasions makes clear that change is expected. Furthermore, McLeod argued that both the social setting in which the questionnaires are completed (i.e. the client's relationship with their therapist and the therapeutic environment, as well as their sense of themselves within that context), and the way that the client makes sense of the questionnaire items, inevitably change over the course of therapy ("the *response shift* phenomenon", p.219), thereby complicating the interpretation of change in scores as an indicator of therapeutic outcome. In particular, he outlined a range of factors that appear to create a response shift, such as *telescoping* (an initial over-estimation of problems), *illusory mental health* (the denial of emotional or psychological difficulties, while at the same time appearing troubled and experiencing physical symptoms), *benign self-deception* (a positively biased, rather than accurate or realistic, view of self), and simply becoming more familiar with the therapeutic model and its terminology. McLeod concluded that the response shift phenomenon is increasingly being understood as a significant methodological problem for the use of self-report instruments in measurement (p.222).

Indeed, Blount, Evans, Birch, Warren and Norton (2002) identified the potential for *paradoxical deterioration* (p.161) when using self-report methods to measure outcome because the process in therapy naturally resulted in a change of perception about problematic experiences and feelings, adding that more thoughtful responses at the end of therapy "may reveal apparent worsening, or

may underestimate improvement, even though more adaptive coping strategies may have been learnt" (p.162). Conversely, Truijens (2017) demonstrated the problems involved in relying only on quantitative self-report data by deconstructing the assumptions underpinning individual symptom measurement using a case study in which the participant's difficulties throughout their period in therapy was not reflected in their scores on the outcome measures used.

Indeed, research into the experience of participants completing self-report questionnaires has identified numerous issues that may affect the way that a participant responds to a self-report questionnaire, thereby threatening the validity of the data being collected. An overview of three examples from the literature reveals the following difficulties for participants completing self-report instruments: the time taken to complete measures, vague, confusing and biased language, irrelevant content, various issues with structure, and upsetting or depressing responses to some of the items (Blount et al., 2002); frustration caused by the instrument's inability to engage with complexity, individuality and meaning (Felton, 2005); and the impact on the client of difficulties in interpreting items in a straightforward way, the limitation of using a prescribed rating scale to represent their experience, and awareness of having an evaluative 'audience' (Truijens et al, 2019).

What are We Measuring?

In order to conduct outcome measurement, whether for the purpose of theory development, evidence-based practice or practice-based evidence, the intended outcome must be defined and a valid instrument that measures that theoretical construct identified. In counselling, as in other social sciences, this is a political decision often influenced by other people's agendas: for example, researchers seeking grant funding designated for research with a particular client group, or a counselling service introducing ROM to demonstrate value-for-money to current or potential funders. Elliott (2002) described this process as "render[ing] unto Caesar", drawing similarities between the taxation of Jews by the Roman Empire, and the relationship between counselling practitioners and researchers,

and the government, NHS, and other bodies with the power to dictate which psychotherapies and services receive investment. He argued that as we depend on Caesar to provide funding for counselling services and research activities, then we have to pay his 'taxes' in the form of quantitative outcome data. The challenge in counselling is whether to define outcome according to medical model concepts (e.g. reduction of symptoms), theoretical models of change (e.g. congruent functioning) or by another metric (e.g. economic impact such as use of NHS services) that may fit the expectations or objectives of the funder. If this decision is made thoughtfully then it is likely to be informed by two additional factors: the heterogeneity of clients, issues and instruments; and the challenge of maintaining a non-diagnostic stance.

Heterogeneity of clients, issues and instruments. Counselling, whether statutory, voluntary or private, tends to be organised as either generalist or specialist services. For example, counselling in the NHS in Scotland - at the current time and where it exists - is offered as a generalist service receiving referrals not referred to alternative services such as clinical psychology or cognitive behavioural therapy. In contrast counselling in England and Wales is offered in the NHS via the *Increasing Access to Psychological Therapies* (IAPT) programme and counselling services are specifically orientated to working with people experiencing depression and anxiety. Voluntary counselling agencies have developed historically to be organised by 'issue' (e.g. bereavement, childhood sexual abuse) because affected individuals have been motivated to develop a specific support service for others. Generalist voluntary counselling services, such as those provided by Crossreach, the social care arm of the Church of Scotland, are far less common. Counsellors in private practice have traditionally offered their services on a generalist basis but, as the market grows, there is a developing trend to identify areas of specialism as a way to differentiate oneself as a practitioner. Nevertheless, clients attracted to services, whether generalist or specialist, are diverse in nature and have idiosyncratic reactions to the experiences that have brought them into counselling. Therefore, even services that offer counselling for people with specific difficulties or

experiences may not necessarily be able to isolate a single construct that is common and meaningful to all clients who access their service. It is a challenge and ethical consideration for any counsellor or counselling service conducting measurement that what they seek to measure is appropriate and relevant to those from whom they gather data.

Non-diagnostic approach. Many instruments have been developed to measure psychological constructs that reflect models of diagnostic classification such as those described in latest edition of the *Diagnostic and Statistical Manual of Mental Disorders* (DSM-5). McLeod (2001) noted the issues implicit in using questionnaires in counselling that reflect a conceptualisation of the person that is different to that of the therapeutic approach. Person-centred counselling takes a non-diagnostic approach to distress (Sanders, 2017), although this stance is under pressure in many contexts. Indeed, a growing number of services use diagnostic measures as means of selecting clients for access to counselling (Evans, 2019). This is counter to the values of counselling and highlights the ethical and political issues involved when making the decision about what to measure, discussed next.

Ethical Issues in Using Measurement in Counselling

The main ethical issues in counselling research are informed consent, minimisation of harm, and confidentiality (Barker et al. 2016). Research studies based on measurement, using aggregated data collected on self-report instruments, are generally considered to have lower risk than studies that adopt case study or qualitative methodologies because confidentiality can more easily be protected via the quantification of responses and anonymisation of participants' identities within large datasets. Indeed, the *Ethical Guidelines for Research in the Counselling Professions* (BACP, 2019) acknowledged this in its advice that, should a participant withdraw consent in quantitative research then, while "it may be impracticable to remove someone from calculations already undertaken […] it may be possible to meet [General Data Protection Regulation] legal requirements by use of anonymisation or pseudonymisation" (p.47). However, the way in which measurement is collected and used in counselling settings raises several ethical

issues, which Evans (2019) summarised as overvaluing the numbers and dehumanising therapy. These issues impact on both clients and counsellors and are exacerbated, I believe, when measurement in counselling is misused or, indeed, abused.

In a recent webinar, Evans (2019), who is one of the main architects of the ROM movement in the UK, argued that it is not appropriate for measurement in counselling to be treated as a high precision science. Instead, it should be viewed as an imperfect method for comparing individuals. He cited various barriers for obtaining valid data from clients using self-report instruments such as cognitive and educational challenges, and language and cultural issues, and raised a critical question about the genuine nature of informed consent for measurement in counselling: do clients who are asked to provide outcome data in counselling believe that they have the right to refuse? Consent to take part in research alongside their counselling when self-referring to a research clinic such as the Strathclyde Counselling and Psychotherapy Research Clinic may appear clear-cut, with appropriate processes and time taken to ensure clients are informed about the potential risks of participating. However, even in this scenario, clients may feel obliged to consent in order to access the counselling offered and, as Stone and Elliott (2011) reported, many find the research process to be hindering in some way. For clients who have accessed counselling within the NHS or a community service conducting ROM, the right to refuse may be far less explicit. Clients may believe that their access to counselling is dependent on taking part in outcome measurement. Indeed, the time taken to complete the instruments and the potential diversion of focus - on to the service's agenda rather than the client's – may significantly detract from the potential of the therapeutic process: it is unlikely in this setting that clients are asked if the selected measures are relevant to their reason for accessing counselling or helpful in their process.

Nevertheless, depending on the purpose for measurement, their scores may have consequences: in particular, when measurement scores are used to classify individuals - and indeed, as Evans' (2019) data demonstrated, misclassify individuals

- when access to a counselling service is dependent on their scores falling within a particular range. This process of inclusion and exclusion based on the reduction of the client's experience to a number highlights the potential for the dehumanisation of therapy.

Indeed, the introduction of IAPT within NHS counselling services in England and Wales provides a case study that demonstrates how measurement within counselling can be abused, with the numbers becoming more important than the people who produce them. IAPT established a target-orientated culture based on efficiency and bureaucracy (Proctor & Hayes, 2017) in which measurement is enforced as a performance management system. Funding of IAPT services depends on the use of a minimum data set (MDS) in which quality is defined as at least 50% of clients achieving 'recovery' according to their post-therapy scores (Delgadillo, Asaria, Ali & Gilbody, 2016). Indeed, as Proctor and Hayes (2017) described, decisions about client progress in IAPT services are more likely to be made by managers, who may not have any counselling training or experience, based on MDS scores, rather than by counsellors in consultation with their clients and supervisors.

Jackson (2019) described the impact of this culture on services and clients: for example, a counsellor interviewed anonymously by Jackson reported that her agency had stopped offering a long-established specialist service for refugees and asylum seekers because these clients' recovery rates were bringing down the agency's overall performance and managers were worried that the agency would lose its funding. This anecdotal report is predicted by findings in the literature. Conducting a study using the national IAPT dataset, Delgadillo et al. (2016) found evidence that services in areas of socioeconomic deprivation had a greater prevalence of mental health problems and lower recovery rates, leading to a significantly larger number of IAPT services in these areas being classified as underperforming. As the researchers stated, this finding poses an ethical and political dilemma for the IAPT programme: should standards be lowered or funding increased for services operating in these deprived areas?

Unsurprisingly, this measurement culture has had a huge impact on counsellors, including stress, burnout and low morale (Jackson, 2019: Surviving Work, 2019). Proctor and Hayes (2017) reported changes to the working conditions of counsellors contracted to IAPT services, including the introduction of payment by results linked to recovery rates defined by clients' MDS scores and the use of recovery rates and targets within services to create competition between individual counsellors. This focus on results is not supported by the theory, practice and research of counselling and this lack of understanding and respect for counsellors and the nature of their work will inevitably have an impact on the quality of the counselling that they can provide for their clients. Thus, rather than measurement providing a reflection of the counselling that takes place, it impedes it. Indeed, the negative impact of the process on the validity of the data generated may be further reinforced because it is typically the counsellor who administers the MDS to the client each session. Therefore, the way in which they introduce it, and handle the process with the client, will have an impact on the client's attitude towards the instrument and the quality of the data collected. Indeed, in a recent survey, 41% of 550 current and past IAPT workers reported that they had been asked to manipulate the performance data collected by their service (Surviving Work, 2019). The most common methods included: coaching clients to give 'good' answers; discharging clients early if scores indicated recovery; and completing questionnaires for clients who had self-discharged.

Given these accounts from the frontline of IAPT, which represents a nationwide implementation of measurement in counselling, it is no wonder that counsellor antipathy toward measurement in counselling has increased. It has confirmed fears that the products of ROM can be used for evaluative, economic and political purposes, and demonstrated the severe consequences - for all - that can result.

So, Why Use Measurement in Counselling?

Given the challenges and ethical issues that are associated with using measurement in counselling, it makes sense that its value can be questioned.

However, I propose that to lose the perspective that an informed approach to measurement offers would significantly restrict our potential to investigate and learn about the experience of counselling: measurement offers a way to approach certain questions that other methods do not. Indeed, to fully embrace the possibilities of research in counselling, methodological pluralism (e.g. Barker et al., 2016) is essential. In contrast with Bondi and Fewell (2016), I propose that both experience-near and experience-far perspectives are necessary, as together they enable us to develop our understanding – to extend our nomological nets - in ways that would be impossible if we had access to only one or other perspective. Measurement allows us to identify potential patterns that may be unrecognisable when counselling or researching with people on a one-by-one basis. It provides a bird's eye view that enables us to identify individuals within our sample whose data appears to match or differ from the perceived pattern, allowing us to approach and explore individual experiences, at ground level, and expand our understanding of the phenomenon.

Therefore, in my view the best way to use measurement within counselling is to understand it as a template that provides an outline, important but limited: the picture only begins to gain meaning when accompanied by the detail and context provided by other perspectives. In other words, theory development in counselling cannot grow without the contribution of *both* measurement *and* qualitative accounts of clients and therapists' experiences in counselling. Similarly, the strengths (and weaknesses) of evidence-based practice can only truly be explored and evaluated when tested in the real world. Finally, the effectiveness of counselling via routine outcome monitoring only makes sense if the meaning of the story told by the data can be understood in the context of the people and the service from which it was collected.

How do we respond to the challenges for measurement in counselling? First, we need to ensure that any measurement conducted in counselling is high quality, based on robust and relevant instruments, with data collected and interpreted in an ethical and appropriate way; second, for this to be the case, significant work is

required so that everyone (e.g. counsellors, researchers, managers, funding bodies) who seeks to use measurement in counselling understands its assumptions, strengths and limitations. This will increase the chances of informed interpretation, better contextualisation and more ethical use of the results, and ultimately improve the relationship between counsellors and measurement, and the validity of the data generated.

Approach to Measurement in this Study

It has been my intention in this study to bridge the two epistemologies described in this chapter in order to conduct a robust investigation of the Strathclyde Inventory for use as an outcome measure in therapy, informed by a methodological pluralism that embraces both the Standards, representing a contemporary understanding of good practice in measure development, and also the inclusion of an experience-near perspective in order to test and understand the meaning of scores collected using the instrument in practice. Therefore, in this study I have gone beyond the typical limits of measure development by engaging with case study data using a mixed method approach.

The Standards

I have approached this study with the understanding, highlighted by the Standards, that validity is not a property of the instrument itself but of the meaning of the scores collected on it. In other words, investigating the validity of scores collected using an instrument being used for specific purposes with a specific population is an ongoing process. In this study I have focused on the potential validity of scores collected on the Strathclyde Inventory from participants representing a UK-based clinical population when interpreted as a measure of outcome of their experience in PCT.

Adopting the Standard's view of validity as a unified concept, I have ensured that my study has gathered evidence from the five recommended sources: test content, response processes, internal structure, related variables, and consideration of potential intended or unintended consequences of the interpretation of scores

for a specific use. The way in which I have organised the data analyses arising from these sources is explained in the next section.

Convinced of the importance of reliability/precision as evidence of fairness, I selected Rasch measurement, a robust model representing the methodology of modern test theory, to interrogate the reliability/precision of my data and to assist in the detection and differentiation of systematic and random errors in measurement.

Slaney and Maraun (2008) proposed a sequential framework to guide and enable researchers to draw on many of the sources of validity evidence outlined by the Standards. They argued that, first, researchers must provide evidence of the internal relationship between the test items and the reliability/precision of scores, and only then can the external validity of those scores (that is, the interpretation of those scores in relation to the theoretical construct that the instrument intends to measure) be meaningfully explored.

As a result, the validation programme presented in this investigation follows the sequence proposed by Slaney and Maraun (2008). In Chapter 4, I have addressed the first two stages in the framework: first, I tested the internal structure of the item responses collected from my dataset by using Rasch measurement to assess the fit of the data to a unidimensional model, and then evaluated evidence of its reliability/precision. In the process, I used the data to generate evidence of validity associated with internal structure, response processes and test content.

In Chapter 5, I began the third stage of the framework: the process of investigating external test validity. I considered various models for testing change in scores within the dataset, including observed, detected and predicted difference, and compared change in scores on the Strathclyde Inventory with that detected by other instruments designed to measure related variables representing outcome. Then I began the process of considering how this information might develop an understanding of the theoretical construct underpinning the Strathclyde Inventory,

congruent functioning, and if and how it appears to change during the course of therapy.

In Chapter 6, I continued my investigation in the form of a mixed methods study that enabled me to gather evidence of external validity from the fifth source proposed by the Standards: the potential consequences of interpreting the pre-post difference in scores on the Strathclyde Inventory as an indication of change during therapy. Would evidence examined at case study level support the classification of clients' outcome from therapy, based on their SI scores, as either improvement or deterioration?

A Mixed Methods Approach to Validation

As Cronbach and Meehl (1955, p.289) observed: "One of the best ways of determining informally what accounts for variability on a test is the observation of the person's process of performance."

Initially, I was inspired by the growing literature supporting the use of mixed methods approaches to validation, most recently reviewed by Zhou (2019). However when I looked more closely at what this meant in practice it became clear that the existing model for using mixed methods in scale development and validation is not a major departure from established practice: an initial qualitative exploration of the theoretical construct, the conversion of qualitative findings into scale items, and quantitative validation of the measure. The main innovation in using mixed methods in validation, as described by Zhou, was a mixed validation approach to reviewing item content based on reflection, debriefing and panel review (qualitative) and sorting and calculation (quantitative) prior to administering the measure for quantitative validation.

Instead, I turned to alternative approaches to outcome within counselling research as potential means for producing evidence that could support or challenge the credibility of using the Strathclyde Inventory as an outcome measure. Outcome studies using qualitative methodology have grown in popularity (e.g. McLeod, 2011), particularly because of its accessibility and resonance for practitioners, and a

process of meta-synthesis, an adapted form of meta-analysis, has developed to aggregate the findings of a group of studies exploring the same themes (e.g. Timulak, 2007; 2009). In addition, systematic approaches to case study research (e.g. McLeod, 2010; Elliott, 2014, 2015) have provided mixed methods models for the integration of quantitative and qualitative data in the evaluation of outcome in individual experiences of therapy.

These perspectives on outcome research enabled me to expand the validation process into the real experience of clients in counselling, thereby introducing an experience-near perspective to the experience-far approach typically employed in measurement studies. This innovation was implemented in my third study, presented in Chapter 6, in which I conducted a two-stage study using, first, case study research methods then qualitative meta-synthesis to compare the outcome for selected participants based on an interpretation of the change in scores on the Strathclyde Inventory, with a much fuller evaluation of the outcome of their therapy based on a range of quantitative and qualitative data collected from the clients and also their therapists.

I began this validation study believing that pursuing a methodological pluralism would significantly enhance the evidence generated in a way that felt entirely appropriate for measure development within counselling. It also seemed important to pursue because, I anticipated, it would provide a more reflective consideration of the potential contribution and relevance of the Strathclyde Inventory for counsellors who may be interested in using this instrument in their work. Finally, it offered me a way to ensure that both my Counsellor self and my Researcher self were fully involved in this investigation – and to discover if it was possible to navigate a path that would challenge and inspire both perspectives.

Chapter 4: An Evaluation of the Internal Structure and Reliability/Precision of the Strathclyde Inventory using the Rasch Model

Chapter Overview

Before investigating the evidence to support the use of the Strathclyde Inventory (SI) as a measure of therapy outcome, it was necessary first to examine its ability to measure the construct that it proposed to measure, using the population whose data were being used for this assessment. Therefore, in this chapter I will focus on the first of my overarching research questions: Are scores on the Strathclyde Inventory valid and reliable measures of congruent functioning when used with a UK-based clinical population? The goal of this study was to go beyond the questions previously explored by Freire (2007) and Zech et al. (2018) by using more robust methods to challenge the internal structure and reliability/precision of the instrument when using data collected from a UK-based clinical population. In essence, this study aimed to assess the evidence that supported the interpretation of a score collected on the SI as a measure of a client's capacity for congruent functioning when completing the instrument *on a single occasion*. If there was sufficient evidence to support the interpretation of scores in this way then it would become possible to investigate the evidence for its use to measure change between scores collected from the same client on two or more occasions. This will be explored in Chapter 5.

Research Questions

This first study was guided by the following questions:

1. Is the SI internally consistent?
2. Does the SI measure a unidimensional concept as proposed by Rogers?
3. Can participants distinguish between the five points in the category structure used by the SI rating scale?
4. Is the SI able to distinguish meaningful levels of *person ability* amongst participants?

5. Are these levels of ability replicated across two different client groups: the heterogenous practice-based protocol and the homogenous social anxiety protocol?
6. Do measurement gaps or redundancies exist in the SI hierarchy of item difficulty that indicate the need to add or remove items?

Method

Study Design

My intention from the beginning was to go beyond the conventional classical test theory (CTT) analyses that have been utilized in previous psychometric studies of the SI (Freire, 2007; Zech et al., 2018). I discovered that advocates of modern test theory (MTT), such as Bond & Fox (2015) and Slaney, Storey & Barnes (2011a, 2011b) have highlighted the advantages for researchers in using MTT methods because they are capable of addressing the traditional concepts of CTT while at the same time providing an expanded ability to test the instrument: broader conceptions of validity and reliability/precision combined with more specific 'localised' estimates to evaluate score precision and validity. I noted Holden & Marjanovic's (2011, p.286) view that, despite these advantages, there has been slow progress in introducing MTT theory and practice into psychology training courses. They argued that it is not surprising that many test developers prefer to continue using familiar CTT procedures. In this study I made the decision to challenge myself, as well as the instrument, by adopting an MTT approach, selecting Rasch measurement, a method underpinned by item response theory, as my chosen model.

Originally developed for application in the field of educational testing, Bond and Fox (2015, pp.50-53) encouraged the wider use of Rasch measurement, advocating it as a useful ideal for the development of practical and robust measures for a broad range of disciplines. The Rasch model is a mathematical framework based on the principles of unidimensionality, estimation (difficulty, ability and precision) and fit (quality control); it rejects CTT assumptions that all items are equally weighted in difficulty and assessed by interval scales (e.g. De Vellis, 2017

p.122). Instead, Rasch measurement transforms raw scores into log odds ratios on a common item-person interval scale (Bond & Fox, 2015, p.364). The unit of measurement is the logit. Test-taker (or *person*) scores are transformed into logits and analysed at individual item-level. The Rasch model assumes that a person whose *ability* matches an item's *difficulty* on the item-person logit scale has a 50% probability of 'success' on that item. If a person's ability is higher on the logit scale than an item's difficulty then this increases the probability of that person achieving a high score on the item (Bond & Fox, 2015, p.46). This process enables *expected scores* to be identified. These are used to detect problems with categories in rating scales (e.g. the assumption that it is an interval scale), the impact of mis-fitting items and persons, and any gaps in the hierarchy of item difficulty that may reduce the accurate measurement of person ability, especially at the higher and lower ends. In addition, organizing expected scores into a hierarchy of item difficulty and person ability can enable researchers to assess if these results are consistent with existing understanding of the theoretical construct being measured or if they introduce a new proposition. Bond & Fox (2015) emphasized that Rasch measurement works best within a thoughtful theory-driven approach to measure development rather than as a routine application of standardized cut-off criteria. It seemed to me that using this method should lead to a clearer understanding of the dimensionality of the theoretical construct measured by the SI but also take me closer to an understanding of the relative difficulty of items for individuals and therefore some evidence of the internal structure of the construct at item-level.

To enable comparisons with existing and future studies that explore the internal structure and consistency of the SI using CTT, I decided to also include the main statistical analyses associated with CTT (e.g. reliability analysis, principal component analysis) and these are reported in relevant sub-sections within the Results section. An analysis of test-retest reliability (temporal consistency) using data from this dataset is presented in Chapter 5.

In a final decision about the design of the study, I chose to include participants in this first study from both protocols conducted in the Strathclyde

Counselling and Psychotherapy Research Clinic ('the research clinic') in order to maximise the size and variability of the potential dataset and to offer an opportunity to include some comparisons between a heterogenous clinical population and a group of participants who were assessed during screening as fitting the diagnostic criteria for a specific psychological difficulty, social anxiety. As I anticipated that participants from the PB protocol may have responded somewhat differently to the items contained in the SI compared to participants from the SA protocol, I have included some analyses in this study in which the data collected from the two sub-samples who accessed the different protocols can be compared.

Participants

Clients. The participants were clients of the research clinic between 2007 and 2016. All clients had consented to take part in research activities alongside the counselling process. The sample contained the data of 385 participants who had accessed either the generalist 'practice-based' protocol, which was non-specific about the nature of the difficulties that participants may wish to work on in therapy and offered up to 40 sessions of PCT with the possibility of extension if mutually agreed (PB; N=294), or a specialist protocol offering 20 sessions of PCT or emotion-focused therapy (Elliott et al., 2004) to people experiencing social anxiety difficulties (SA; N=91). The majority of participants in both protocols were female (PB – 65.3%; SA – 56%) and white European (PB – 96.7%; SA – 94.1%) and were of a similar mean age (PB – 35.9 years, range 18-73; SA – 33.2 years, range 18-60). The two research protocols were originally approved by the NHS ethics committee (PB: 7 May 2008; SA: 29 April 2008) and the university's ethics committee (UEC; PB: 30 August 2007; SA: 14 May 2007). Ethical approval of the research clinic's generic framework, which includes within its scope all three studies that form this thesis, was most recently renewed by the UEC on 10 April 2018 (see Appendix B).

Researchers. The data were collected by students and staff who participated as researchers in the research clinic between 2007 and 2016. I contributed to the data collection process for the SA protocol as a volunteer from 2008-2011 then to

the PB protocol as research clinic coordinator from 2013 onwards. Professor Robert Elliott was the principal investigator of both protocols.

Data Collection

Participants completed the SI at regular time points during their involvement with the research clinic: before therapy began, at designated mid-points (in the PB protocol, following every 10th session of therapy; in the SA protocol, following the 8th session of therapy), at the end of therapy, and at two follow-up points (six and 18 months after the end of therapy). Thus, the number of data collection points at which the SI was completed per participant depended on the number of counselling sessions they accessed. In this dataset the range was 1-10 (M=2.89; SD = 1.72; median = 3; mode = 1) offering a full dataset of 1174 observations.

Of these, 652 observations were collected using Version 4 of the SI (Elliott & Rodgers, 2007; PB = 436; SA = 216) then, from 2012 onward, the remaining 522 observations (PB = 410; SA = 112) were collected using Version 6 (Elliott & Rodgers, 2012c). For the remainder of this study these versions will be named as SI-31 (Version 4) and SI-16 (Version 6). The two versions of the SI shared sixteen items; those items that had been retained from SI-31 by Elliott & Rodgers (2012b) when creating SI-16 (Appendix C). In order to create a dataset that included all observations, irrespective of version, I extracted the data collected on SI-31 for these 16 items and created a dataset in which all observations were represented by scores obtained on these 16 items (numbered below according to SI-16):

1. I have been able to be spontaneous
2. I have condemned myself for my attitudes or behavior (R)
3. I have tried to be what others think I should be (R)
4. I have trusted my own reactions to situations
5. I have experienced very satisfying personal relationships
6. I have felt afraid of my emotional reactions (R)
7. I have looked to others for approval or disapproval (R)
8. I have expressed myself in my own unique way
9. I have found myself “on guard” when relating with others (R)

10.I have made choices based on my own internal sense of what is right
11.I have listened sensitively to myself
12.I have lived fully in each new moment
13.I have hidden some elements of myself behind a “mask” (R)
14.I have felt true to myself
15.I have been able to resolve conflicts within myself
16.I have felt it is all right to be the kind of person I am

Items marked (R) indicate those negatively worded items that require to be reverse scored. As it is possible that participants completing these sixteen items in the form of SI-16 may have responded somewhat differently to participants who encountered the items embedded amongst others in the form of SI-31, I have included some preliminary analyses in which the data collected from the two sub-samples who completed different versions of the SI-16 can be compared.

As a dataset containing multiple observations obtained from the same participants was likely to violate the key statistical assumption of independence of observations (Field, 2013, p.176), I decided to create a subsample that included only one observation per participant. Although the simplest approach would have been to work with pre-therapy observations only, I wanted to include data collected from participants at later stages in their therapeutic process, as I hoped that this would broaden the dataset: I expected participants who were in therapy to have developed in relation to the construct being measured. Therefore, I produced a sub-sample that deliberately but randomly included observations collected at a range of time points across therapy. To achieve this, I used an online true random number service (random.org) to select one observation from each participant who had completed the instrument on more than one occasion. This process resulted in an ‘independent’ dataset of 385 observations. Table 4.1 presents the original full dataset and the subsequent independent dataset, identifying sub-samples from each protocol and each version of the SI.

Table 4.1. Study 1 datasets presented by protocol and SI version

		Total	PB		SA	
			SI-31	SI-16	SI-31	SI-16
Full dataset	N	1174	436	410	216	112
	%		37.1%	34.9%	18.4%	9.5%
Independent	N	385	158	136	58	33
dataset	%		41.0%	35.3%	15.1%	8.6%

Notes. PB = practice-based protocol; SA = social anxiety protocol. SI-31 = Strathclyde Inventory (31 item version); SI-16 = Strathclyde Inventory (16 item version).

As Table 4.1 shows, there was little change in the proportion of each version of the SI represented when the independent dataset was created: SI-31 = 55.5% (full); 56.1% (independent). However, there was some change observed in the proportion of each protocol present in the reduced dataset. As can be seen, the independent dataset contained a larger proportion of observations from participants in the PB protocol than the full dataset. This indicates that a larger proportion of participants in the SA protocol contributed multiple observations to the full dataset. This trend can be seen in Table 4.2, in particular when comparing percentages at the 1st session, Mid-1, post-therapy and two follow up data collection points.

Table 4.2. Study 1 datasets presented by protocol and data collection point.

	Full dataset (N=1174)				Independent dataset (N=385)			
	PB		SA		PB		SA	
Total N	846	%	328	%	294	%	91	%
Pre-therapy	302[a]	35.7	96[a]	29.3	168	57.1	36	39.6
1st session	65	7.7	46	14.0	13	4.4	8	8.8
Mid-1	132	15.6	66	20.1	33	11.2	18	19.8
Mid-2	91	10.8	0	0	19	6.5	0	0
Mid-3	61	7.2	5	1.5	10	3.4	0	0
Mid-4	15	1.8	3	0.9	3	1.0	0	0
Mid-5	5	0.6	2	0.6	0	0	0	0
Mid-6	1	0.1	0	0	0	0	0	0
Post-therapy	106	12.5	61	18.6	28	9.5	17	18.7
Follow up (6 months)	46	5.4	32	9.8	10	3.4	9	9.9
Follow up (18 months)	22	2.6	17	5.2	10	3.4	3	3.3

Notes. PB = practice-based protocol; SA = social anxiety protocol. [a] total includes second observations completed by some clients before therapy commenced.

It is possible that this reflects more successful data collection processes achieved because of the shorter duration of therapy and more robust administrative practices implemented by the principal investigator and volunteers operating the SA protocol.

Table 4.2 presents a detailed comparison of the data collection points represented by the observations collected in the full and independent datasets. Interestingly, the percentage of observations obtained at pre-therapy increased as a proportion of the independent dataset. This can be explained easily: this was the only data collection point for those participants with only one observation in the dataset and was also an option within the randomization process. Therefore, although the process to create an independent sample removed potential issues arising from non-independence, the possibility of results from this sample being influenced by the dominant presence of pre-therapy scores increased. Nonetheless, I had succeeded in creating an independent sub-sample that included representations of the construct as experienced by participants across the therapeutic process.

Other measures. The SI was administered to participants by their researcher as the second or third instrument within a battery of outcome measures. The two other instruments that formed the standard set of outcome measures for both protocols were CORE-OM (Evans et al., 2002; Connell et al., 2007), designed to measure general concepts of distress and functioning, and the Personal Questionnaire (PQ; Elliott et al., 2016), an individualized instrument created by the client at the intake interview in which they itemised specific difficulties that they wished to address in therapy. I will present more detailed information about these two instruments in Chapter 5 as I conducted analyses that compare and correlate change in scores on the SI, PQ and CORE-OM as part of my second study.

Participants in the SA protocol completed additional outcome measures targeting specific issues associated with social anxiety difficulties. None of the data collected on these other instruments are included in this thesis.

Period of measurement and rating scale. When completing the SI, participants were asked to read each statement, consider how often it had been true for them during the last month, then mark the box that was closest to their experience using a 5-category rating scale with the following anchor words: never, only occasionally, sometimes, often, always.

Data Analysis

I used SPSS software (versions 23-25) to test internal consistency and structure according to CTT and Winsteps (version 3.62.1; Linacre, 2006) for the analyses using Rasch measurement.

Results

The results of my analyses are reported in the following order: internal consistency, dimensionality, precision/reliability of the 5-category rating scale, evidence of meaningful and distinct levels of person ability that are consistent across different client groups, and consideration of gaps and redundancies in the hierarchy of item difficulty.

Is the SI Internally Consistent?

Internal consistency and correlation. First, I used CTT techniques to explore the consistency of the SI's internal structure. Internal consistency is a measure of the homogeneity of items within an instrument; in other words, the degree to which the items selected are measuring the same variable. It is typically reported using Cronbach's coefficient alpha (α), the proportion of total variance in the data that can be explained by the common variance (covariance) among the items (De Vellis, 2017, pp.44-43). Inter-item covariance represents the variance reported as raw scores; inter-item correlation is a standardized version. In Table 4.3, I have presented the Cronbach's alpha and mean inter-item correlation for the independent dataset as a whole, then by sub-sample (protocol and SI version).

Table 4.3. Internal consistency and inter-item correlation of independent dataset by protocol and SI version

	Total	Protocol		SI version	
		PB	SA	SI-31	SI-16
N	374[a]	286[a]	88[a]	209[a]	165[a]
Cronbach's alpha	.92	.92	.87	.92	.91
Inter-item correlation (mean)	.41	.43	.30	.42	.39

Notes. [a] = SPSS has deleted cases listwise where variables missing.

As Table 4.3 shows, the SI demonstrates strong internal consistency based on the data collected for this study. A Cronbach's alpha for the whole sample of .92 is very high; with little variation across sub-samples. The mean inter-item correlation (*r* = .41) is also acceptable, with only the data of one sub-sample, the SA participants, indicating a slightly lower degree of correspondence between items.

While in general a higher alpha is better as it provides evidence of good reliability, Barker et al. (2016, p.70) have argued that internal consistency greater than .90 may be a sign of too many items (overkill) or that the variable being measured may be too easily rateable (triviality). This has been a consistent result for the SI across different length of versions and groups of test-takers (see Chapter 2) and suggests that the instrument could be shortened further. The slightly lower Cronbach's alpha and mean inter-item correlation produced by the data collected from the SA participants provides some evidence of greater variability in rating, thereby disputing the potential triviality of the variable, as perceived by participants. Later analyses in this chapter using Rasch will also enable the manner in which participants respond to the measure to be investigated.

Rasch analysis. In Rasch measurement, the starting point for assessing internal consistency and precision/reliability is an examination of the standardised residuals for person and items, known as the fit statistics. These indicate the degree to which the instrument, in this case the SI-16, fits the Rasch model. The results demonstrate that the overall fit found between persons, items and the Rasch model was excellent. Table 4.4 presents a range of fit statistics, first for persons, the

sample of participants (N=385) and then for items, the collection of items that make up the instrument (N=16).

Table 4.4. SI-16 fit statistics for persons & items.

					Infit		Outfit	
	Score	Count	Measure	Error	MNSQ	ZSTD	MNSQ	ZSTD
Persons (N=385)								
Mean	30.3	16.0	-.11	.34	1.01	-.2	1.01	-.2
S.D.	11.7	.3	1.07	.08	.60	1.6	.58	1.6
Real RMSE	.35	Adj. SD	1.01	Separation	2.92	Person Reliability		.90
Items (N=16)								
Mean	728.3	383.1	.00	.06	1.00	-.3	1.01	-.2
S.D.	113.3	.8	.42	.00	.26	3.8	.26	3.7
Real RMSE	.06	Adj. SD	.42	Separation	6.48	Item Reliability		.98

Notes. MNSQ – mean square; ZSTD = standardized mean square; SD = standard deviation; RMSE = real root mean square error; Adj. SD = adjusted standard deviation

For persons, the mean raw score on the SI-16 was 30.3 (SD = 11.7), demonstrating that there was a substantial variation in raw scores between persons. The statistic labelled 'count' represents the number of items for which persons returned scores. 'Measure' is the person's (or item's) raw score transformed into logits, the Rasch unit of measurement. For persons, the measure represents their 'ability', in this case their ability in terms of congruent functioning, according to the instrument. Therefore, the mean measure (-.11) represents the mean ability of all persons in the sample. A perfect fit between person ability and item difficulty according to the Rasch model would be zero, therefore this result shows that the match is very close.

This result is demonstrated again in the infit and outfit statistics for persons. These are summarized statistics expressed both as mean squares and in a standardized form. The infit statistic is a weighted residual placing more emphasis on unexpected responses close to a person's (or item's) measure. The outfit statistic is unweighted, produced by averaging the residual variance across persons and items; therefore, it can be influenced by unexpected responses that are far from a person's (or item's) measure (Bond & Fox, 2015, p.355). The infit and outfit

statistics for persons matched: 1.01 (*Z* = -.2). The predicted mean square values for infit and outfit statistics to indicate a good fit to the Rasch model is 1; the expected value for *Z* is 0 and the recommended cut-off is +/-2.0. (Bond & Fox, 2015, p.356). However, Bond & Fox (2015, p.274) also noted that larger sample sizes are more likely to lead to mean square fit statistics that are close to 1.0 and that the smallest amount of misfit can register as significant according to *Z.* Therefore, they recommend using this data as evidence to be weighed with other test information as well as the researcher's familiarity with the test context. In this case, the mean squares indicate a good fit but the standardised statistics, both negative, suggest that there is a small likelihood of less variation in person ability than expected in the model, a potential 'overfit' to the Rasch model that may result in smaller standard errors and inflated separation and reliability statistics, but also may have no practical implications (Bond & Fox, 2015, p.270-1).

Certainly, the person reliability statistic was found to be high at .90, only slightly lower than the Cronbach's alpha (.92), which is based on the same concept and reported in Table 4.3. According to Bond & Fox (2015, p.49), high person reliability typically indicates there is a large enough spread of ability across the sample to produce consistent patterns of persons with higher and lower scores. The person separation statistic, derived from the real RSME and adjusted standard deviation statistics, was 2.92. I will explain and discuss this result in more detail later in this chapter.

I found similar results in the item fit statistics. In general, these indicated that the 16 items were a good fit to the Rasch model. Again, the mean score (728.3; SD = 113.3) shows that there was some variation in the data. The mean count was 383.1, revealing that there was some missing data given that the total number of persons in the sample was 385. The mean measure for item difficulty (0.0) is expected as this is routinely set at zero in Rasch measurement (Bond & Fox, 2015, p.365). There was only tiny variation in the infit and outfit mean-square statistics (infit = 1.00; outfit = 1.01) but some difference in the standardized statistics (infit *Z* = -.3; outfit *Z* = -.2). As before, this negative infit *Z*, plus the exceptionally high item

reliability statistic (.98), suggests the likelihood that there may be some degree of overfit amongst the items, most likely that they lack local independence; in other words, that some may be too similar or redundant.

To identify where some of this overlap may reside, I examined the misfit statistics for each individual item. These are presented in Table 4.5 and, replicating the output produced by Winsteps, are ordered by outfit mean square.

Table 4.5. Item statistics: misfit order.

Item	Measure	Model S.E.	Infit		Outfit		Point Measure Correlation
			MNSQ	ZSTD	MNSQ	ZSTD	
7	.26	.06	1.42	5.5	1.51	6.4	.54
6	-.34	.06	1.49	6.2	1.45	5.8	.55
5	.12	.06	1.43	5.6	1.40	5.2	.58
9	.52	.06	1.09	1.2	1.10	1.4	.63
1	.17	.06	1.05	.7	1.09	1.2	.58
3	-.20	.06	1.07	1.1	1.07	1.1	.66
10	-1.07	.06	1.06	.9	1.06	.8	.57
13	.32	.06	1.03	.5	1.02	.3	.67
2	.13	.06	.96	-.6	.97	-.3	.65
11	-.05	.06	.87	-2.0	.94	-.9	.68
16	-.18	.06	.85	-2.3	.85	-2.2	.73
8	-.34	.06	.83	-2.7	.83	-2.6	.68
12	.80	.06	.81	-2.9	.79	-3.1	.72
4	-.31	.06	.77	-3.6	.76	-3.7	.69
14	-.18	.06	.64	-6.1	.63	-6.1	.77
15	.32	.06	.59	-6.9	.61	-6.6	.77

Notes. S.E. = Standard Error; MNSQ – Mean Square; ZSTD = T statistic; S.D. = Standard Deviation

The outfit mean squares for all individual items range from 1.51 (item 7 – *I have looked to others for approval and disapproval*) to .61 (item 15 – *I have been able to resolve conflicts within myself*). These results are within the range (0.5 – 1.5) identified as productive of measurement by Linacre (2017, p.349) and acceptably close to the range (0.6 – 1.4) recommended by Bond & Fox (2015, p.273) for scores obtained from rating scales. The point-measure correlation indicates how well the individual item aligns with the SI-16 as a whole; Table 4.5 demonstrates that all items are consistent with the overall measure.

The *Z* scores reflect the ordering of items by mean square values but contain some unexpectedly large results: nine items have infit and outfit *Z* scores that exceed +/-2.0. Three of these items (7 - *I have looked to others for approval and disapproval*; 6 - *I have felt afraid of my emotional reactions*; 5 - *I have experienced very satisfying personal relationships*) have positive *Z* scores, indicating underfit: 'noisy', unpredictable, unexpected responses, possibly the result of 'poor' items or participants having 'special' knowledge or guessing. The negative *Z* scores reported for the remaining 6 items (15 - *I have been able to resolve conflicts within myself*; 14 - *I have felt true to myself*; 4 - *I have trusted my own reactions to situations*; 12 - *I have lived fully in each new moment*; 8 - *I have expressed myself in my own unique way*; 16 - *I have felt it is all right to be the kind of person I am*) suggest overfit: 'too good to be true', likely due to item dependence (Bond & Fox, 2015, p.272). The best fitting items are items 13 (*I have hidden some elements of myself behind a 'mask'*), 10 (*I have made choices based on my own internal sense of what is right*), 3 (*I have tried to be what others think I should be*), 1 (*I have been able to be spontaneous*), 9 (*I have found myself 'on guard' when relating with others*), 2 (*I have condemned myself for my attitudes or behaviour*) and 11 (*I have listened sensitively to myself*).

As reported earlier, it is possible that the large sample size explains these standardised statistics, picking up small misfits and amplifying them. However, to conclude this might be a missed opportunity to consider the potentially useful information that they provide about the items. Taken as a whole, these results suggest that, according to the Rasch model, some items may compromise the integrity of the SI-16 because the patterns of participants' scores are either too random or too predictable. There will be more opportunities to investigate the performance of these items as this study unfolds.

Does the SI Measure a Unidimensional Concept?

Principal component analysis (CTT). I conducted a principal component analysis with varimax rotation (PCA) in order to report the internal structure of the SI-16 from a CTT perspective. This was the method used by previous researchers (Freire, 2007; Zech et al., 2018) and therefore the most appropriate choice because

the purpose of the analysis was to enable comparison of the data collected from my dataset with that of previous studies. As Fokkema & Greiff (2017, p.401) noted, PCA is similar but not the same as exploratory factor analysis (EFA) although it is often identified as such in the literature. They argued that PCA is a formative measurement model in which principal components represent a "parsimonious summary" of the item scores; in other words, a model that assumes that the item scores cause the construct. In contrast, EFA is a reflective measurement model that assumes that it is the construct that causes the item scores.

First, I conducted the PCA with the whole sample, then on sub-samples by protocol and version. I checked the Kaiser-Meyer-Olkin measure of sampling adequacy for each analysis. Following my initial analysis, in which I did not limit the number of components that could be extracted, I repeated the analysis but restricted to one component to see how much variance could be explained by a one-solution model using this method. The results of these analyses are presented in Table 4.6.

Table 4.6. Principal component analysis (CTT) of independent dataset by protocol and SI version

	Total	Protocol		SI version	
		PB	SA	SI-31	SI-16
Observations in sample (N)	385	294	91	216	169
KMO measure of sampling adequacy	.93	.94	.81	.93	.90
Total variance explained (%)	54.3	56.6	67.3	55.9	59.8
Number of components extracted	2	2	5	2	3
Component 1 (rotated):					
Variance explained (%)	31.7	32.3	22.0	32.7	27.3
N items	10	10	8	10	9
Component 2 (rotated):					
Variance explained (%)	22.6	24.3	15.5	23.2	18.4
N items	6	6	3	6	4
Component 3 (rotated):					
Variance explained (%)	-	-	12.6	-	14.1
N items			2 [a]		3[e]
Component 4 (rotated):					
Variance explained (%)	-	-	10.4	-	-
N items			2 [b]		

Component 5 (rotated):					
Variance explained (%)	-	-	6.9	-	-
N items			1 [c]		
% variance explained by 1 component	44.9	47.3	35.2	46.7	43.0
N items loading <.4 on 1 component	0	0	1 [d]	0	0

Notes. [a] = items 4 & 10; [b] = items 2 & 6; [c] = item 7; [d] = item 6; [e] = items 5, 9 & 13.

In contrast, the item scores contained in the two smallest sub-samples (SA protocol and SI-16 version) suggest a five component and three component solution, respectively, with a higher degree of variance explained by these models compared to the other whole and sub-samples, in particular for the SA protocol sub-sample (67.3%). However, none of the three additional components suggested by the SA protocol data contained a sufficient number of items - three or more, according to Gorsuch (1997) - to be considered meaningful. For the item scores collected using SI-16, there is one additional component and this does contain three items: items 5 (*I have experienced very satisfying personal relationships*), 9 (*I have found myself 'on guard' when relating with others*) and 13 (*I have hidden some elements of myself behind a 'mask'*). It is unclear why these items would have a different relationship on the shorter version of the instrument. It may be that participants' perception of these items changed with the removal of other items. Alternatively, it may be that these variations are more visible within the smaller sub-samples. Certainly, when absorbed into the whole sample, the influence of these somewhat different scoring patterns is not visible. It is worth noting here that there appears to be no obvious relationship between these item clusters and the misfitting items noted earlier in this chapter; the clusters contain a mixture of apparently overfitting, underfitting and well-fitting items.

Finally, on reanalyzing the data but restricting the number of components to one, I found that the amount of variance explained by the model fell across all samples but particularly the two smallest sub-samples: whole sample (44.9%), PB protocol (47.3%), SA protocol (35.2%), SI-31 version (46.7%) and SI-16 version (43.1%). There was only one item in one sub-sample (SA protocol) that loaded <.4

on a one-component model; this was item 6 (*I have felt afraid of my emotional reactions*), one of the underfitting items identified earlier.

Principal component analysis (Rasch). Rasch measurement promotes the investigation of unidimensionality of the measured construct by conducting its own form of principal components analysis (PCA). Instead of basing the investigation of structure on correlations of sample-dependent ordinal data as CTT factor analyses do, the Rasch PCA, having transformed the raw scores into logits that can be measured on a linear model, examines the standardized residuals that remain after the linear Rasch measure has been extracted, looking for indications of any other common variance within the residuals (Bond & Fox, 2015, p.284-5). Table 4.7 presents the standardized residual variance, calculated as eigenvalue units.

Table 4.7. Standardized residual variance in SI-16

	Eigenvalue Units	% of Total Variance	% of Unexplained Variance
Total variance in observations	40.8	100.0%	-
Variance explained by measure	24.8	60.7%	-
Unexplained variance (total)	16.0	39.3%	100%
Unexplained variance in 1st contrast	2.5	6.2%	16.2%
Unexplained variance in 2nd contrast	1.7	4.1%	10.2%
Unexplained variance in 3rd contrast	1.4	3.5%	8.6%
Unexplained variance in 4th contrast	1.3	3.2%	8.2%
Unexplained variance in 5th contrast	1.1	2.8%	6.5%

The total variance found in the data measured 40.8 eigenvalue units. Of this, 24.8 units (60.7%) was explained by the measure, greater than the typical value (40% - 50%) for explained variance that supports the presumption of unidimensionality according to Linacre (2017, p.555). This result is also substantially higher than the amount of variance explained by a one-component solution when tested using the CTT form of PCA.

A central feature of Rasch principal components analysis is the ability to contrast the extracted residuals. To support the proposition of unidimensionality, residual variance should be random and without structure (Linacre, 2017). In my analysis, five contrasts were extracted, but only the first contrast was greater than

2.0 eigenvalues, the proposed lower size limit for random variance warranting further investigation recommended by Linacre (2017, p.555).

Table 4.8. Item loadings in first contrast

Plot Point	Item	Loading
A	3 - I have tried to be what others think I should be (R)	.62
B	2 - I have condemned myself for my attitudes or behavior (R)	.54
C	6 - I have felt afraid of my emotional reactions (R)	.53
D	7 - I have looked to others for approval and disapproval (R)	.52
E	9 - I have found myself “on guard” when relating with others (R)	.29
F	13 - I have hidden some elements of myself behind a “mask” (R)	.29
G	5 - I have experienced very satisfying personal relationships	-.23
H	4 - I have trusted my own reactions to situations	-.24
h	16 - I have felt it is all right to be the kind of person that I am	-.24
g	10 - I have made choices based on my own internal sense of what is right	-.27
f	14 - I have felt true to myself	-.29
e	15 - I have been able to resolve conflicts within myself	-.32
d	12 - I have lived fully in each new moment	-.41
c	1 - I have been able to be spontaneous	-.42
b	11 - I have listened sensitively to myself	-.42
a	8 - I have expressed myself in my own unique way	-.44

Table 4.8 presents the loadings of items onto the first contrast, while Figure 4.1 presents these data in the form of a plot that clearly reveals a pattern within the contrast: one cluster of positively loaded items (the six reverse scored items) and one cluster of negatively loaded items (the ten other items).

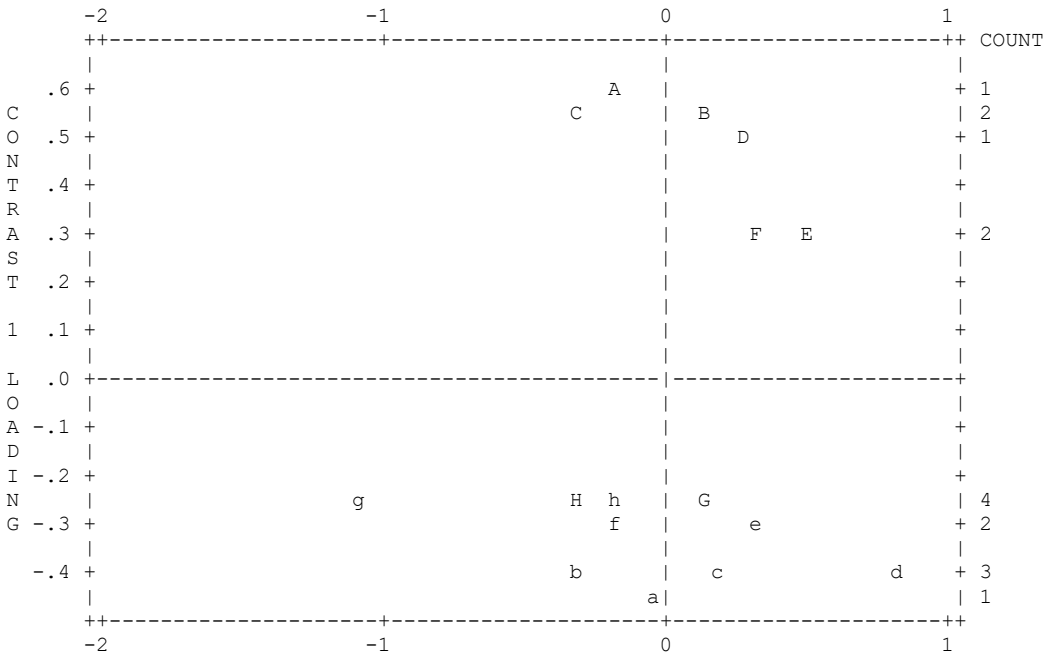


Figure 4.1. Plot of items in first contrast
Note. Refer to Table 4.8 to identify items that match plot A-H and a-h.

As recommended by Linacre (2017, pp.556-7), I calculated mean person measures for the two sub-groups, first, positively loaded items, then, negatively loaded items, and correlated the results. The Pearson correlation was .64, suggesting that the positively and negatively loading items in the first contrast are not measuring something substantially different. While it seems that there is some pattern that can be detected in the way that participants approach these two groups of items, the most likely explanation seems to be the way in which these groups of items are worded, positively or negatively. I propose that this result confirms that this first contrast (which corresponds to the second component identified in the first PCA; Table 4.6) does not represent a second dimension measured by the SI but is a common artefactual factor produced by the effect of two groups of items being worded in the opposite direction caused by some form of acquiescence response from participants (e.g. Gorsuch, 1997; Spector, Van Katwyk, Brannick & Chen, 1997; De Vellis, 2017 pp.117-118).

Can Participants Distinguish Between the Five Points in the Category Structure Used by the SI Rating Scale?

It is fairly typical to find that participants have had difficulties in using a five-point rating scale because of varying interpretations of the meaning of anchor words, as Elliott et al (2006) noted. Rasch measurement is able to examine the functioning of the category structure underpinning an instrument's rating scale. It does this by calculating infit and outfit mean squares for each response category and threshold calibrations that indicate the measure in logits at which the most probable rating option, based on item difficulty and person ability, changes from one category to another. All person and item data can be used in this analysis as Rasch subtracts item measure from person measure and, based on this information, calculates the most probable responses across a scale that represents the range of difference in the dataset. Typically, this analysis is conducted at the beginning of a study using Rasch measurement as, if problems with the category structure are detected, then these can be resolved before moving on to other analyses. Following this approach, I did conduct this analysis as a first step in this study but as a result of the results, presented below, no remediation work was required. I have presented the results of this analysis at this point in the Results section because it seemed to me that it was a better fit for the flow in my presentation.

Table 4.9 presents the category structure revealed by this analysis.

Table 4.9. Summary of the SI-16 rating scale category structure

Category Label	Score	Count	Infit MNSQ	Outfit MNSQ	Threshold Calibration
Never	0	709	.94	.95	None
Only occasionally	1	1662	.92	.91	-1.86
Sometimes	2	1878	.93	.93	-.51
Often	3	1286	.94	.95	.61
All or most of the time	4	594	1.28	1.31	1.76

Note. MNSQ = Mean Square

Table 4.9 shows the total frequency (count) for each category and the infit and outfit mean squares for each category, all which fall well within the acceptable

range (0.6 – 1.4) proposed by Linacre & Wright (1994). This means that the participants' responses on the rating scale were not more nor less predictable than can be tolerated by the Rasch model. The category that contained most random responses (infit mean square = 1.28; outfit mean square = 1.31) was 4, 'all or most of the time' (or 'never' for items that were reverse scored), indicating that there was approximately 30% more randomness in participants' choice of this category than predicted by the model. This was also the category used least frequently, consistent with its position as an extreme category within the 5-point scale. It seems probable that a small number of outliers may have influenced the fit statistics for this category by selecting a higher (or lower, if reversed) score on certain items in a way that was contrary to the majority of participants with an otherwise similar level of ability. Possible explanations for these unpredicted responses might include social desirability (e.g. to be seen as doing well), lack of self-awareness, or misunderstanding the meaning of the item.

It can be seen in Table 4.9 that the thresholds between response categories increased monotonically across the measure. This indicates that, on average, persons with increasing ability endorsed progressively higher categories. Linacre (1999) has recommended that there should be at least 1.0 logit between thresholds in a 5-category scale, which was the case here. The ordered nature of the SI-16 rating scale can be seen clearly in the form of a probability curve, as demonstrated in Figure 4.2.

As Figure 4.2 shows, the thresholds between categories – the points at which the most probable response moves into the next category within the rating scale - are distinct, although the peaks for categories 2 and 3 are slightly lower than 50% probability. Indeed, this probability curve is unusually clean and clear, indicating that participants were able to distinguish between the five points on the rating scale, leading to the conclusion that the SI-16 rating scale was functioning as intended.

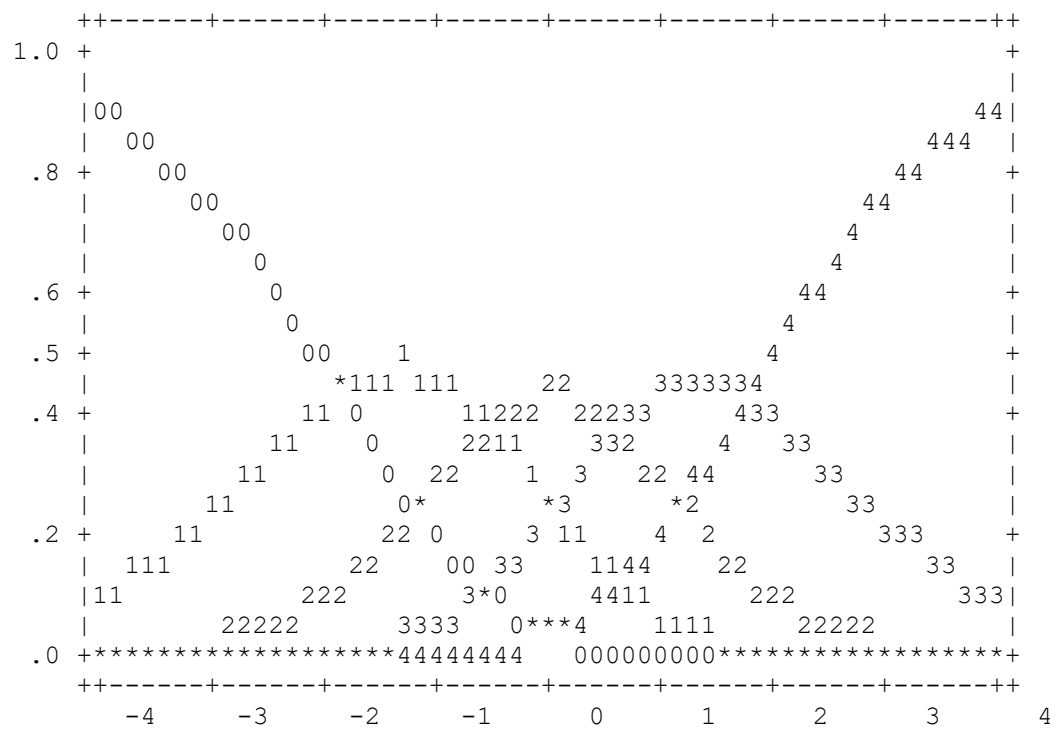


Figure 4.2. Category probability curve of the SI-16 rating scale
Notes. x-axis = person measure minus item measure (logits); y-axis = probability of response (percentage).

Is the SI Able to Distinguish Meaningful Levels of 'Ability' Amongst Participants?

The initial Rasch analysis carried out to produce standardised residuals for persons and items identified a person separation index (*G*) of 2.92 (see Table 4.4). This metric can be converted into a strata index using the formula [(4*G*-1)/3] provided by Bond & Fox (2015, p.354); a strata index represents the number of measurably distinct strata (i.e. layers or levels) of person ability (or item difficulty) that can be supported by the data. This calculation for persons was 3.6, indicating that the SI-16 can distinguish at least three distinct levels of 'ability' in the data. This suggests that the SI-16 is capable of measuring individuals in different strata and therefore that it should be possible to detect movement of persons between strata; in other words, that it is appropriate for use as an outcome measure. Using the same formula, the item separation index of 6.48 converted to 8.3 strata, indicating at least eight steps in the hierarchy of item difficulty contained in the SI-16.

These item and person strata are visible in Figure 4.3, an 'expected score' item-person matrix showing the average category rating predicted for each item from a person with a specific measure of ability.

To understand Figure 4.3, it should be noted that the scale for person ability is on the x-axis, while items are displayed on the y-axis in vertical order of increasing difficulty. The expected category ratings for each item are presented horizontally across the matrix, increasing as person ability increases. The half-point threshold between categories on each item is represented by a colon.

Below the matrix, the distribution of persons is shown: M indicates the mean of the person measures; S indicates one standard deviation on either side of the mean; T indicates two standard deviations on either side of the mean. As highlighted earlier in Table 4.4, the average person is almost in line with the mid-point (-.11) of the x-axis, indicating an overall good match between the sample and the measure. Quite simply, the three levels of ability proposed by the person strata could be understood as locating in the zone below -S (low ability), the zone between -S and +S (average ability) and the zone above +S (high ability).

Items are ordered within the matrix ranging from those that participants found easiest to endorse, therefore requiring least ability (item 10 – *I have made choices based on my own internal sense of what is right*), to those found to be most difficult to endorse, therefore requiring most ability (item 12 – *I have lived fully in each new moment*). An image that can be helpful in making sense of this item-person matrix is to see it as a ladder in which each strata, which could be an individual item or a cluster of items, is a rung that marks progress in a person's development of ability in relation to the attribute that the ladder measures. A perfectly constructed ladder will have just enough rungs with just enough distance between them to enable each person to climb as far up the ladder as their current ability allows. If the SI-16 is well designed then this ladder will be an accurate representation of the theoretical construct it intends to measure: congruent functioning.

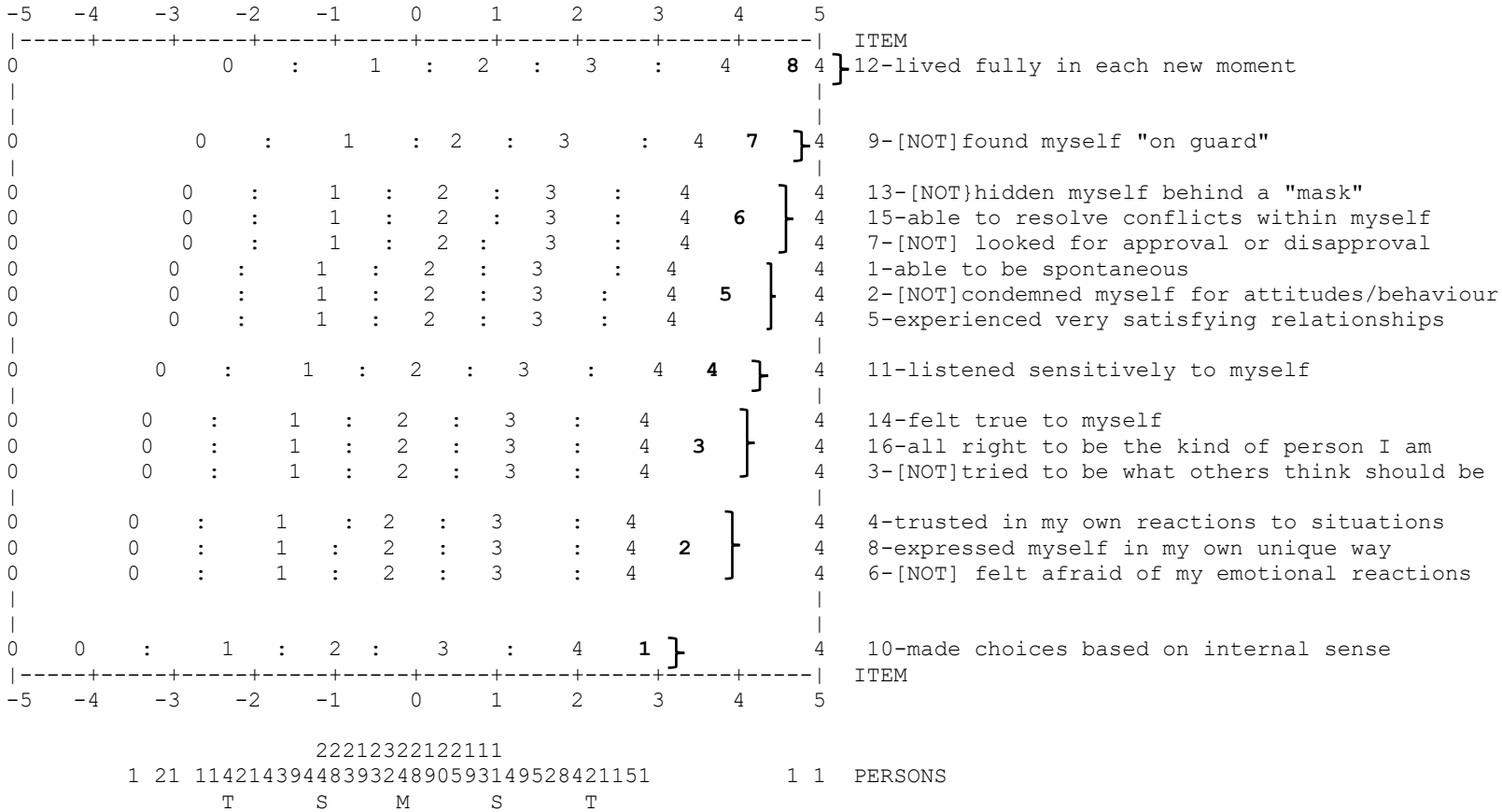


Figure 4.3. SI-12: Expected score item-person matrix

Notes. “0,1,2,3,4” represents the mean expected score category selected by person according to measure on x-axis; ":" indicates Rasch-half-point threshold; numbers below x-axis are total number of persons (presented vertically) at each measure point; M= mean person measure; S = one standard deviation from mean; T = two standard deviations from mean; bracketed item or cluster of items indicates proposed item strata.

Therefore, my next step was to review the items within each strata to see if the sequence of rungs in the SI-16 ladder make sense from a theoretical perspective.

The first strata consisted of item 10 (*I have made choices based on my own internal sense of what is right*) only, suggesting that most participants considered that, in the last month, they had the capacity to make decisions based on their ability to sense the right choice for them. To me, this did not fit theoretically as a first step. Based on my understanding of the construct, I would expect that someone with a lower level of congruent functioning is more likely struggle to know and act on their own sense of what is right. Something else would have to occur first: increased awareness of what they felt and believed and a growing willingness to trust in themselves. So, this first rung in the ladder felt a little wobbly to me, as if it was out of place or had taken the first position due to a missing rung.

Moving on to the second strata, I found a cluster comprising items 6 (*I have felt afraid of my emotional reactions*; a reverse scored item), 8 (*I have expressed myself in my own unique way*) and 4 (*I have trusted my own reactions to situations*). Together, these items suggested that those participants who were able to score higher on these items were able to trust and appreciate their reactions and responses as they occurred. This seemed to fit better as an early part of the process of developing congruence. The third level of ability is indicated by another cluster of three: items 3 *(I have tried to be what others think I should be* – reverse scored), 16 (*I have felt it is all right to be the kind of person I am*) and 14 (*I have felt true to myself*). This strata seemed to reflect a growing recognition and acceptance amongst participants of the value of being themselves. Theoretically this made sense as a development of the experience of being able to trust, understand and appreciate their reactions as indicated by the previous rung.

The fourth strata comprised a single item (11 - *I have listened sensitively to myself*). This suggested that those participants who have come to value themselves can then also recognize the importance of listening carefully to themselves. Again, this seems like a logical progression, theoretically. Next came another three-item cluster: items 5 (*I have experienced very satisfying personal relationships*), 2 (*I have*

condemned myself for my attitudes or behavior; reverse scored) and 1 (*I have been able to be spontaneous*). The items that formed this fifth strata suggested a sense of relief and release: being able to enjoy the people in one's life, being less harsh on oneself and feeling able to act impulsively. This implied a growing trust in, and valuing of, self that is being acted upon. The sixth strata contained items 7 (*I have looked to others for approval or disapproval*; reverse scored), 15 (*I have been able to resolve conflicts within myself*), and 13 (*I have hidden some elements of myself behind a 'mask'*; reverse scored). Participants who were able to score on these items seemed to share a growing confidence in themselves, particularly in their relationships with other people: less reliant on others for their sense of worth, more willing to reveal themselves, and a greater sense of their own competence to understand and resolve contradictory needs that they experience.

The final two strata, or rungs on the SI-16 ladder, were both represented by individual items. Item 9 (*I have found myself 'on guard' when relating with others*; reverse scored) formed the seventh strata, suggesting that, even at this late stage, some participants, while able to take off their 'mask', had nevertheless held onto some degree of guardedness as a protection in certain relationships and that, at last, this becomes unnecessary due to an increased ability to trust and accept themselves *and* others. The freedom associated with this experience can be seen clearly in the eighth strata, item 12 (*I have lived fully in each new moment*). Participants were able to recognize the experience of being fully themselves, without barriers (masks or guards). It fits theoretically that this item represented the zenith of ability, according to the scores of participants in this dataset.

In summary, this hierarchy of item difficulty suggests that the SI-16 can measure growing layers of development – self-awareness, self-trust, self-acceptance, leading to greater openness to self and to others – that are consistent with the process of congruent functioning. However, it did seem to me as if the lowest rung of the ladder did not quite fit its function. I will look at this issue later in this chapter. Before that I present what I discovered when I separated out the two sub-samples within my dataset – PB protocol and SA protocol – to find out if this

hierarchy of item difficulty held when comparing these differently constituted groups of participants.

Are these Levels of Ability Replicated across both Client Groups: The Heterogenous Practice-based Protocol and the Homogenous Social Anxiety Protocol)?

In order to test if the levels of ability identified for the whole dataset were replicated across two different client groups, one heterogeneous in relation to their reasons for accessing therapy, the other more likely to be homogenous, I carried out the same analysis separately with each of the two protocol sub-samples. First, I obtained the fit statistics for each sub-sample so that I could identify any noticeable variation.

As Table 4.10 shows, there is very little variation in the fit statistics displayed for each sub-sample when compared with the dataset as a whole and between sub-samples. In particular, the measure for each sub-sample remains close to zero, with the PB sub-sample slightly closer (-.04) than the whole dataset (-.11) and the SA sub-sample somewhat further away (-.38).

The infit and outfit statistics (mean squares and *z* scores) are almost exactly the same in each analysis, with one small but interesting difference: the outfit *z* score for items scored by the PB sub-sample is slightly lower (3.2) than the same statistic for the whole dataset and the SA sub-sample (3.7). This is interesting because the PB sub-sample (N=294) is much larger than the SA sub-sample (N=91). Therefore, it may not be possible to rely on sample size as an explanation of this difference. Instead, this result suggests that there are more misfitting scores affecting item fit within the SA sub-sample and, as a result, the dataset as a whole. This finding is supported by the slightly lower person reliability (.85) and item reliability (.90) statistics identified for the SA sub-sample.

Table 4.10. Standardised residuals for persons & items in whole dataset and two sub-samples

					Infit		Outfit	
	Score	Count	Measure	Error	MNSQ	ZSTD	MNSQ	ZSTD
Persons (N=385; Whole dataset)								
Mean	30.3	16.0	-.11	.34	1.01	-.2	1.01	-.2
S.D.	11.7	.3	1.07	.08	.60	1.6	.58	1.6
Real RMSE	.35	Adj. SD	1.01	Separation	2.92	Person Reliability		.90
Items (N=16)								
Mean	728.3	383.1	.00	.06	1.00	-.3	1.01	-.2
S.D.	113.3	.8	.42	.00	.26	3.8	.26	3.7
Real RMSE	.06	Adj. SD	.42	Separation	6.48	Item Reliability		.98
Persons (N=294; PB protocol)								
Mean	31.0	16.0	-.04	.34	1.01	-.2	1.00	-.2
S.D.	12.2	.3	1.11	.09	.61	1.6	.59	1.6
Real RMSE	.35	Adj. SD	1.05	Separation	3.02	Person Reliability		.90
Items (N=16)								
Mean	568.4	292.3	.00	.07	1.00	-.3	1.00	-.2
S.D.	88.6	.7	.43	.01	.26	3.8	.26	3.2
Real RMSE	.07	Adj. SD	.43	Separation	5.79	Item Reliability		.97
Persons (N=91; SA protocol)								
Mean	28.1	16.0	-.38	.33	1.02	-.2	1.01	-.2
S.D.	9.8	.2	.89	.07	.57	1.6	.56	1.6
Real RMSE	.34	Adj. SD	.82	Separation	2.43	Person Reliability		.85
Items (N=16)								
Mean	159.8	90.8	.00	.06	1.00	-.3	1.01	-.2
S.D.	26.3	.4	.42	.00	.26	3.8	.26	3.7
Real RMSE	.13	Adj. SD	.39	Separation	2.96	Item Reliability		.90

Notes. As for Table 4.4.

The main difference between the three sets of fit statistics can be seen in the person separation and item separation indices. As I explained earlier, these statistics are calculated by dividing the adjusted standard deviation by the real root mean square error. The person separation for the whole dataset was 2.92; for the PB sub-sample it is 3.02, and for the SA sub-sample it is 2.43. The smaller adjusted

standard deviation (.82) for the SA sub-sample is the reason for this difference. This means that there is less spread in the person measures in this sub-sample compared to the PB sub-sample and, as a result, the whole dataset. When recalculated as strata using the formula [(4*G*-1)/3], the difference is small but significant: PB sub-sample = 3.7 and SA sub-sample = 2.9 compared to the result for the whole dataset (3.56). It means that it may not be possible to identify three levels of ability if the SA sub-sample data were analysed on their own.

This lack of differentiation in the SA sub-sample is seen again when calculating strata using the item separation index for each analysis. The item separation indices were: 6.48 (whole dataset), 5.79 (PB sub-sample) and 2.96 (SA sub-sample). Using the same formula with this data results in the following strata estimations: 8.3 (whole dataset), 7.4 (PB sub-sample) and 3.6 (SA sub-sample).

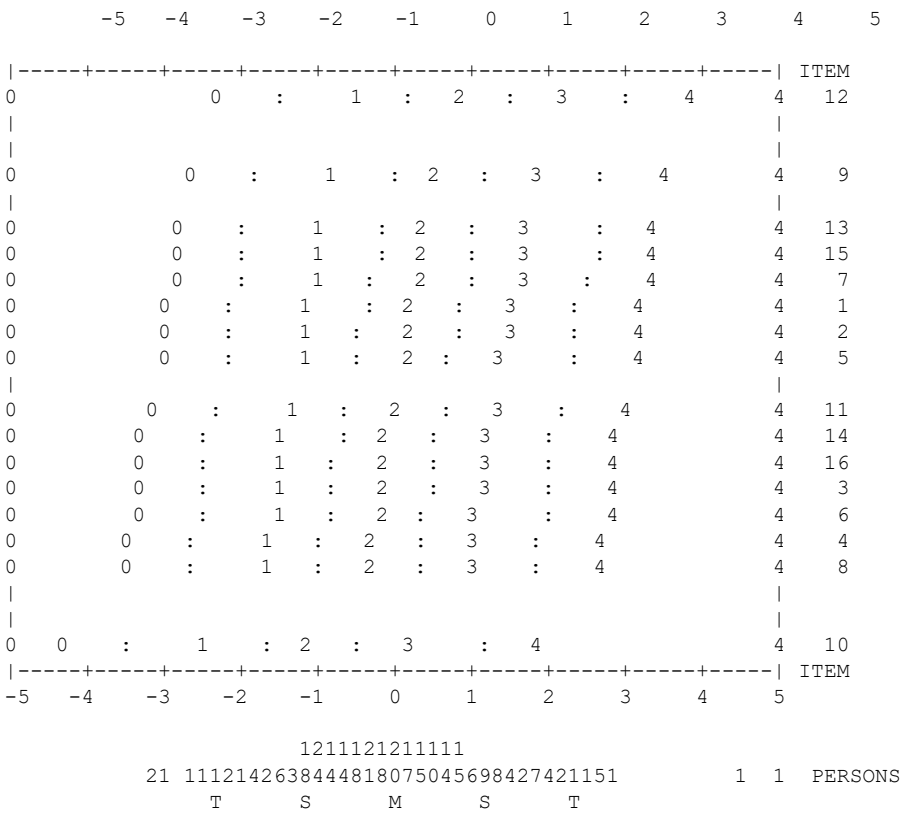


Figure 4.4. SI-16 Expected score item-person matrix (PB protocol only)
Notes. As for Figure 4.3.

This series of findings creates an expectation that the SA sub-sample contains more influential outlying scores (persons and items) than the PB sub-sample and will

produce a narrower, less well-defined expected score item-person matrix when compared with the PB sub-sample and the whole dataset, into which its idiosyncrasies have been absorbed. This suggests a likelihood that there will be visible differences in the three hierarchies of items that the data would produce.

The expected score item-person matrix for each sub-sample is presented as Figure 4.4 (PB protocol) and Figure 4.5 (SA protocol). Note the narrow spread of persons below the x-axis in Figure 4.5.

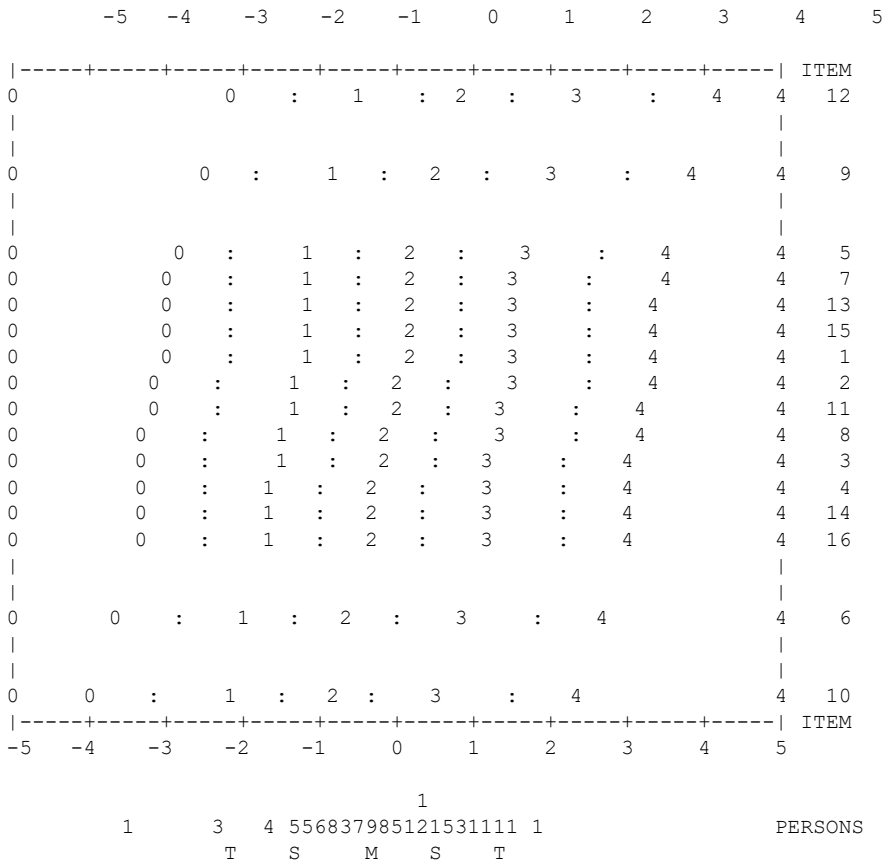


Figure 4.5. SI-16 Expected score item- person matrix (SA protocol only)
Notes. As for Figure 4.3.

To assist in the identification and interpretation of differences between these three matrices (Figures 4.3, 4.4 and 4.5), I have created Table 4.11. This shows the position of items in the hierarchy of item difficulty identified according to Figure 4.3, then in subsequent columns, the hierarchy evident in Figures 4.4 and 4.5.

Table 4.11. Position of items in hierarchy of item difficulty by dataset

Whole dataset		PB	SA	Items (ordered according to SA hierarchy)
Rung	Item	Item	Item	
8	**12**	**12**	**12**	**I have lived fully in each new moment**
7	**9**	**9**	**9** [c]	**I have [not] found myself "on guard" when relating with others**
6	13	13	5 [a]	I have experienced very satisfying personal relationships
6	15	15	7 [c]	I have [not] looked to others for approval and disapproval
6	7	7	13	I have [not] hidden some elements of myself behind a "mask"
5	1	1	15 [b]	I have been able to resolve conflicts within myself
5	2	2	1	I have been able to be spontaneous
5	5	5	2	I have [not] condemned myself for my attitudes or behaviour
4	**11**	**11**	**11** [c]	**I have listened sensitively to myself**
3	14	14	8 [ac]	I have expressed myself in my own unique way
3	16	16	3	I have [not] tried to be what others think I should be
3	3	3	4 [c]	I have trusted my own reactions to situations
2	4	6	14 [b]	I have felt true to myself
2	8	4	16 [bc]	I have felt it is all right to be the kind of person that I am
2	6	8	6	I have [not] felt afraid of my emotional reactions
1	**10**	**10**	**10**	**I have made choices based on my own internal sense of what is right**

Notes. **Bold** = no change in position across all samples. [a] = item has moved into a higher rung on hierarchy (requires more ability for this sub-sample); [b] = item has moved into a lower rung on hierarchy (requires less ability for this sub-sample); [c] = scoring on item disordered according to observed average measures for persons.

As Table 4.11 demonstrates, the hierarchy of item difficulty across the whole dataset and sub-samples is framed at the top, bottom and in the middle by four items, each representing a standalone strata in the ladder, that do not vary: item 12 (*I have lived fully in each new moment*), item 9 (*I have found myself 'on guard' when relating to others*; reverse scored), item 11 (*I have listened sensitively to myself*), and item 10 (*I have made choices based on my own internal sense of what is right*).

Looking more closely, the hierarchy of item difficulty for the PB sub-sample almost replicates that for the whole dataset. The only difference is that item 6 (*I have felt true to myself*) appears to be somewhat more difficult for this group of participants; it moves two places up the hierarchy, but still remains within the cluster identified as the second strata.

As predicted, there is much more variation in the hierarchy of item difficulty produced using scores from the SA sub-sample. Two items moved to a higher strata, appearing to require more ability for these participants: item 5 (*I have experienced very satisfying personal relationships*) and item 8 (*I have expressed myself in my own unique way*). This makes sense as these participants have come to therapy to work with social anxiety difficulties, which these two items seem to reflect. In particular, item 5 jumps to become the third most difficult item in the hierarchy. It was one of the underfitting items identified earlier in this chapter and the substantial change in difficulty provides further evidence that this may be a less useful item for the SI-16 as it has the potential to produce erratic responses.

There are also three items that moved down one strata, appearing to require less ability for these participants: item 15 (*I have been able to resolve conflicts within myself*), item 14 (*I have felt true to myself*) and item 16 (*I have felt it is all right to be the kind of person that I am*). These changes are more difficult to explain based on item content: all seem relevant to the participants' social anxiety difficulties. However, it may be that, in the context of their difficulties as a whole, these items represented less challenging aspects of their lives. These three items were identified as overfitting items earlier in this chapter and this puzzling movement may suggest them for consideration as redundant items.

Finally, I noted when viewing an observed average measures for person matrix produced in the same Winsteps output as the expected score item-person matrix, that the scoring by the SA sub-sample on six items was disordered, in other words that the actual responses chosen, averaged across persons with the same measure (level of ability) did not increase across the matrix in a Guttman pattern of progression (i.e. 0, 1, 2, 3, 4; Bond & Fox, 2015, p.28). This was not the case for the PB sub-sample, and was not visible when looking at the same matrix for the dataset as a whole. The items affected were: item 9 (*I have found myself "on guard" when relating to others*; response order 0,1,2,4,3); item 7 *(I have looked to others for approval or disapproval*; response order 0,4,1,2,3); item 11 (*I have listened sensitively to myself*; response order 0,1,2,4,3); item 8 (*I have expressed myself in*

my own unique way; response order 0,1,2,4,3); item 4 (*I have trusted my own reactions to situations*; response order 0,1,2,4,3) and item 16 (*I have felt it is all right to be the kind of person I am*; response order 0,1,2,4,3). This indicates that sufficient participants in the sub-sample were scoring a '4' on these items when this would not have been predicted by the overall measure of their ability (as demonstrated in the earlier category structure analysis). This pattern seems to have been particularly extreme for item 7 (*I have looked to others for approval or disapproval*) when the average response that followed '0', as ability increased, was '4', demonstrating that some participants with lower overall ability found this item somewhat easier to answer than other participants with higher ability. Once again, item 7 was identified as an underfitting item and this finding supports this conclusion.

This detailed examination of the impact on the hierarchy of item difficulty when the two sub-samples that formed the dataset in this study are separated provides further evidence about the internal structure and precision/reliability of the SI. It has demonstrated that the main framework (top, bottom, middle) for the hierarchy holds across samples. The heterogeneity present in the PB sub-sample appears to have been a significant influence on the results for the whole dataset, which makes sense given its relative size. The impact of the SA sub-sample has been a greater challenge. My series of analyses highlighted that this was a narrow, apparently less well differentiated sample that contained some influential misfitting scores, resulting in a degree of variation from the original hierarchy of item difficulty (Figure 4.3) and that arising from the PB sub-sample (Figure 4.4). Table 4.11 showed that five items moved between strata according to the scores provided by the SA sub-sample. However, this movement was limited to the next strata (e.g. item 5 moved from rung 5 to rung 6; item 14 moved from rung 3 to rung 2) and may suggest that these two strata (2 and 3; 5 and 6) might be measuring a similar stage in the process of development. If so, then there may be a case for removing items that may be redundant within these broad strata. This issue will be considered in the second half of the next section.

What Measurement Gaps or Redundancies Exist along the SI Hierarchy of Items that Indicate the Need to Insert or Remove Items?

Measurement gaps. Having conducted this series of analyses, I felt that there were gaps at the lower end of the hierarchy of item difficulty. In particular, it seemed to me that item 10 (*I have made choices based on my own internal sense of what is right*) did not fit theoretically as a relevant measure of a low level of congruence. As an instrument for use within a clinical setting it seemed important to ensure that the SI was capable of more detailed measurement at the lower end of the construct. Therefore, I reviewed the 15 items that had been extracted from the original 31-item version by Elliott & Rogers (2012b) and selected four items that were removed in a preliminary cull of items with loadings <0.5 without reviewing content. I considered that these four items described experiences potentially of relevance for 'low ability' persons. These items were inserted into a new version (SI-20; Appendix D) in the same position as they had held in the SI-31. The four new items, with their SI-20 numbering, were:

12. I have felt myself doing things that were out of my control (R)
14. I have been aware of my feelings
18. I have felt myself doing things that are out of character for me (R)
19. I have accepted my feelings

Next, I extracted data collected on the 20 items from the SI-31 subsample within the independent dataset (n=216) and repeated the Rasch analyses previously carried out with the SI-16 data. I compared the results of these analyses with the SI-16 findings and found them to be consistent. Table 4.12 presents a comparison of the fit statistics for SI-16 and SI-20.

As Table 4.12 illustrates, the fit statistics for the SI-20 indicates a very similar, if slightly better, fit to the Rasch model. The person measure (-.04) is even closer to zero; the infit and outfit *Z*-scores for the standard deviation of items are slightly lower. I calculated strata for persons to be 3.9 (SI-16 = 3.6) and for items, 8.5 (SI-16 = 8.3).

Table 4.12. Comparison of SI-16 and SI-20 fit statistics for persons & items

	Score	Count	Measure	Error	Infit MNSQ	Infit ZSTD	Outfit MNSQ	Outfit ZSTD
SI-16								
Persons (N=385)								
Mean	30.3	16.0	-.11	.34	1.01	-.2	1.01	-.2
S.D.	11.7	.3	1.07	.08	.60	1.6	.58	1.6
Real RMSE	.35	Adj. SD	1.01	Separation	2.92	Person Reliability		.90
Items (N=16)								
Mean	728.3	383.1	.00	.06	1.00	-.3	1.01	-.2
S.D.	113.3	.8	.42	.00	.26	3.8	.26	3.7
Real RMSE	.06	Adj. SD	.42	Separation	6.48	Item Reliability		.98
SI-20 (revised version)								
Persons (N=216)								
Mean	40.2	19.9	-.04	.29	1.01	-.2	1.00	-.2
S.D.	14.2	.6	1.01	.07	.56	1.6	.52	1.6
Real RMSE	.30	Adj. SD	.97	Separation	3.18	Person Reliability		.91
Items (N=20)								
Mean	431.9	214.0	.00	.09	1.00	-.3	1.00	-.2
S.D.	87.2	1.0	.58	.01	.27	3.1	.27	3.0
Real RMSE	.09	Adj. SD	.57	Separation	6.62	Person Reliability		.98

Notes. As for Table 4.4.

I examined the misfit statistics for the items in SI-20, looking for any visible changes when compared to the statistics for the SI-16 items and checking the results for the four new items. There were some minor shifts in order for a few existing items and a slight reduction in *Z*-scores, probably due to the smaller sample size, but no changes that raised concern. The statistics for the four new items demonstrated a comfortable fit within the parameters set by the existing group of items (see Table 4.13). Item 12 was found to be the second most underfitting item of the group of twenty, but still within acceptable limits: infit = 1.45 (z = 4.3); 1.41 (z = 3.8). Items 18 (infit = 1.29 (z = 2.9); 1.21 (z = 1.9)) and 14 (infit = 1.14 (z = 1.5); 1.16 (z = 1.6)) were better fits. Item 19 was found to be somewhat overfitting but not extreme within the overall group: infit = .79 (z = -2.5); .78 (z = -2.5).

Table 4.13. SI-20 Items: misfit statistics

Item	Measure	Model S.E.	Infit		Outfit		Point Measure Correlation
			MNSQ	ZSTD	MNSQ	ZSTD	
7	.48	.08	1.39	3.9	1.49	4.7	.52
12	**-.83**	**.08**	**1.45**	**4.3**	**1.41**	**3.8**	**.52**
5	.35	.08	1.38	3.8	1.36	3.6	.59
6	-.15	.08	1.37	3.7	1.33	3.3	.57
18	**-1.20**	**.09**	**1.29**	**2.8**	**1.21**	**1.9**	**.53**
10	-.90	.08	1.14	1.5	1.20	2.0	.54
14	**-1.03**	**.08**	**1.14**	**1.5**	**1.16**	**1.6**	**.42**
15(13)	.46	.08	1.13	1.4	1.12	1.3	.63
1	.29	.08	1.08	.9	1.12	1.3	.55
9	.75	.08	1.02	.3	1.01	.1	.64
3	-.04	.08	.97	-.3	.96	-.4	.68
2	.41	.08	.90	-1.1	.93	-.8	.68
8	-.14	.08	.84	-1.8	.83	-1.9	.67
11	.15	.08	.81	-2.2	.84	-.1.8	.68
20(16)	.09	.08	.79	-2.4	.80	-2.3	.75
19	**.01**	**.08**	**.79**	**-2.5**	**.78**	**-2.5**	**.70**
13(12)	1.01	.08	.73	-3.2	.73	-3.1	.72
4	-.10	.08	.67	-4.1	.66	-4.2	.71
17(15)	.42	.08	.59	-5.3	.60	-5.0	.76
16(14)	-.01	.08	.54	-6.0	.55	-5.9	.78

Notes. As for Table 4.5. New items highlighted in bold.

In addition, I checked and identified no meaningful difference in category structure nor in the dimensionality of the instrument.

All but one of the four new items had measures (see Table 4.13) that indicated, as had been hoped, that they would provide a better fit with the needs of lower ability persons. To confirm this observation, I examined the position that the four new items adopted within the expected score item-person matrix, presented in Figure 4.6.

As predicted, three of these items – item 18 (*I have felt myself doing things that are out of character for me*), item 14 (*I have been aware of my feelings*) and item 12 (*I have felt myself doing things that were out of my control*) clustered at the low end of the matrix, below and just above the original lowest difficulty item, (item

10; *I have made choices based on my own internal sense of what is right*). The position of these items made much more sense theoretically, offering additional opportunities to capture the early stages of development of persons at the lower end of the scale and improving the overall validity of the instrument. The fourth addition, item 19 (*I have accepted my feelings*) appeared to have less to offer. It was located in the middle of the hierarchy of item difficulty between items 20 (*I have felt it is all right to be the kind of person I am*) and 16 (*I have felt true to myself*). The similarity in meaning of these three items is reflected in the overfit indicated by item 19's misfit statistics and provided further evidence of item redundancy in the middle of the hierarchy. As Figure 4.6 shows, the strata evident in the hierarchy of items for SI-20 indicates seven strata, rather than eight. There is now more differentiation at the lower end of the hierarchy and less differentiation in the middle, evident in the merging of former SI-16 strata 2, 3, and 5 into SI-20 strata 4, and former SI-16 strata 4 and 5 into SI-20 strata 5.

This finding replicates the strata model indicated by the SA sub-sample data supporting the view that the items that have formed these strata have a tendency to overlap and could be 'thinned out'.

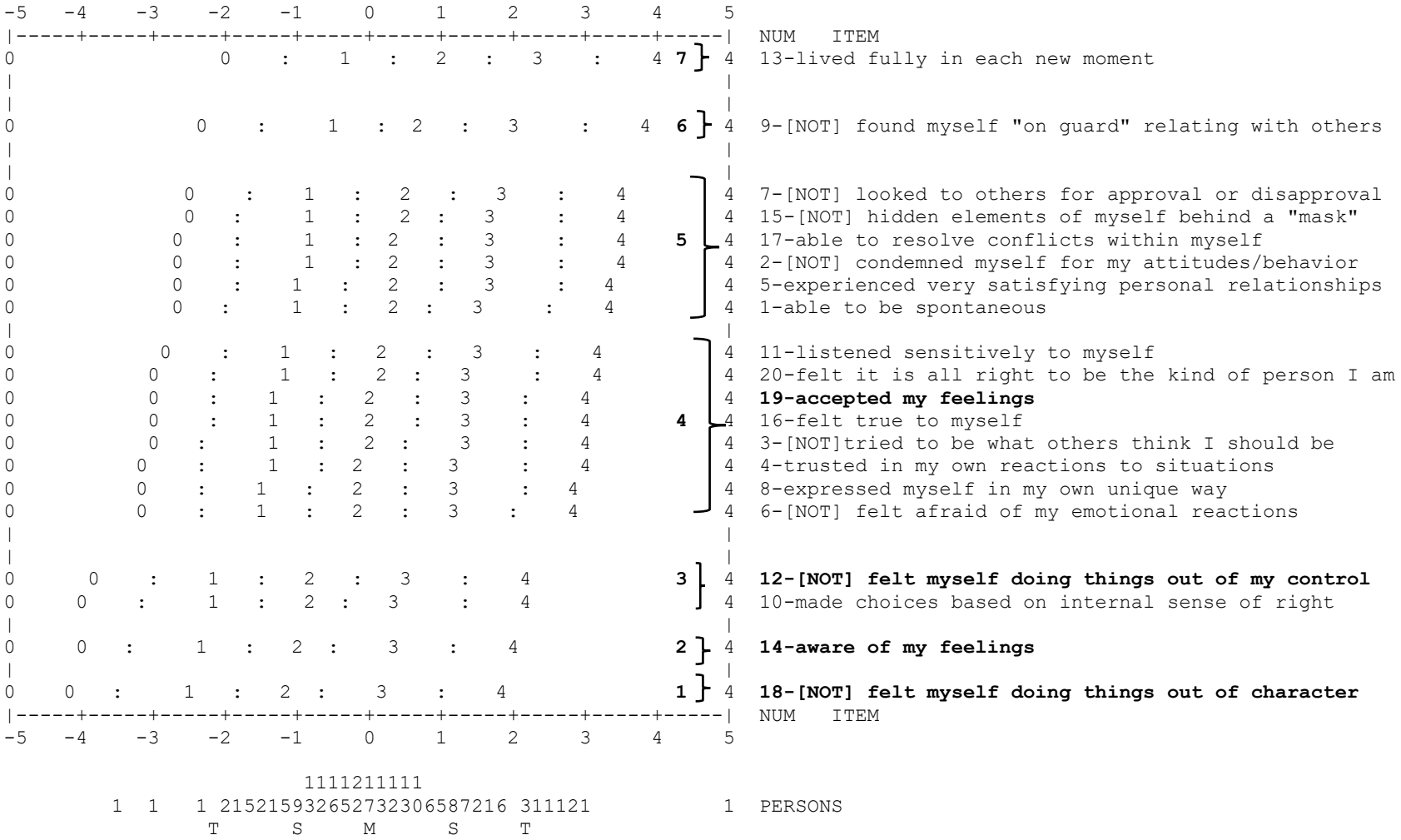


Figure 4.6. SI-20 Expected score item- person matrix

Notes. As for Figure 4.3; in bold = items added to create SI-20.

Item redundancies. Therefore, while re-introducing three items from SI-31 clearly improved the validity of the Strathclyde Inventory at the lower end of the range of difficulty and did not harm its overall fit to the Rasch model, it accentuated the overlap of items in the middle range of difficulty within the instrument. Furthermore, enlarging the instrument was not a desirable outcome; briefer instruments are widely encouraged as they are perceived to have greater fairness in practice because they are less onerous for participants (Rolstad, Adler & Rydén, 2011). Therefore, it was evident that there was a final stage to complete: the identification of items that could be removed from the instrument without harming its demonstrated fit to the Rasch model. Nevertheless, removing items in order to *reduce* person reliability appeared to be a less typical aim amongst Rasch researchers. Advice offered by Linacre (2010) on the removal of items identified the likely main purpose as the *improvement* of person measurement, represented by the person reliability index. However, in this case, the series of Rasch analyses had provided a range of evidence that the number of items could afford to be reduced (high person reliability and the presence of several somewhat overfitting items) and also information about the performance and fit of the items themselves that could guide the decision-making process. My challenge was to identify items that, when removed, would slightly 'loosen' without harming the instrument's overall fit to the model. I delayed conducting this final stage of this study until I had further item level data arising from my second study (details reported in Chapter 5).

I focused my search for redundant items on SI-20 strata 3, 4 and 5 because the other strata, below and above, were each formed by one item only and, although by no means supported by perfect fit statistics especially at the lower end of the scale, were necessary to extend the range of person ability that the instrument could capture.

I gathered together the results of the various item-level analyses that I had conducted and assembled them in Table 4.14.

Table 4.14. Overview of item properties in SI-20 strata 3, 4 and 5 to inform removal process

Item	Measure	Misfit[1]	PMC	*d*	DIF	MVT	DIS	CI-TC	SMC
Strata 5									
7. I have looked to others for approval or disapproval (R)	**.48**	**U**	**.52**	**L**	**N**	**N**	**Y***	**.47**	**.46**
15. I have hidden some elements of myself behind a "mask" (R)	.46	G	.63	S	N	N	N	.57	.48
17. I have been able to resolve conflicts within myself	**.42**	**O**	**.76**	**L**	**N**	**Y**	**N**	**.74**	**.67**
2. I have condemned myself for my attitudes or behaviour (R)	.41	G	.68	L	N	N	N	.64	.55
5. I have experienced very satisfying personal relationships	**.35**	**U**	**.59**	**S**	**N**	**Y**	**N**	**.54**	**.38**
1. I have been able to be spontaneous	.29	G	.55	M	N	N	N	.50	.45
Strata 4									
11. I have listened sensitively to myself	.15	O	.68	M	N	N	Y	.66	.61
20. I have felt it is all right to be the kind of person I am	.09	O	.75	M	N	Y	Y	.78	.66
19. I have accepted my feelings	**.01**	**O**	**.70**	**M**	**N**	**-**	**-**	**.67**	**.59**
16. I have felt true to myself	**-.01**	**O**	**.78**	**N**	**N**	**Y**	**N**	**.76**	**.64**
3. I have tried to be what others think I should be (R)	-.04	G	.68	M	N	N	N	.64	.58
4. I have trusted my own reactions to situations	-.10	O	.71	M	N	N	Y	.68	.54
8. I have expressed myself in my own unique way	**-.14**	**G**	**.67**	**M**	**N**	**Y**	**Y**	**.65**	**.54**
6. I have felt afraid of my emotional reactions (R)	**-.15**	**U**	**.57**	**M**	**N**	**N**	**N**	**.50**	**.44**
Strata 3									
12. I have felt myself doing things that were out of my control (R)	-.83	U	.52	M	N	-	-	.45	.49
10. I have made choices based on my own internal sense of what is right	**-.90**	**G**	**.54**	**S**	**Y**	**N**	**N**	**.54**	**.43**

Notes. **Bold** = items removed. [1] = misfit according to infit *z*-scores: U = underfit ($z > 2.0$); G = good fit ($-2.0 > z < 2.0$); O = overfit ($z < -2.0$). PMC = point mean correlation; *d* = effect size of pre-post change (Chapter 5 results): L = large effect (> .75); M = medium effect (> .45); S = small effect (< .44); N = negative effect (< 0). DIF = differential item functioning (pre-post data; Chapter 5); Y = evidence that item function changed between pre- and post-therapy. MVT = item moved strata in SA sub-sample analysis. DIS = disordered scoring in SA sub-sample analysis; * = considerable disorder. CI-TC = corrected item-total correlation. SMC = squared multiple correlation; Y = SQM > .9.

The data that I included were: measure (representing each item's level of difficulty in SI-20); misfit (the infit *z*-score for the item within SI-20, coded as underfit (z > 2.0); good fit (-2.0 > z < 2.0); overfit (z < -2.0); point mean correlation (PMC; the degree to which the item correlates with the SI-20 as a whole); the effect size of pre-post change on the item (results presented in Chapter 5) coded as L = large effect (> .75); M = medium effect (> .45); S = small effect (< .44); N = negative effect (< 0); the results of a differential item functioning analysis comparing pre-therapy data with post-therapy data (results presented in Chapter 5) where Y indicates that evidence was found that the functioning of the item changed between pre- and post-therapy; and, from the analysis of the data by protocol, if the item moved into another strata or if scoring of the item was disordered according to the SA sub-sample analysis of SI-16.

In addition, I conducted a reliability analysis of SI-20 using SPSS and noted the corrected item-total correlation for each item (corresponding to the Rasch point mean correlation), the squared multiple correlation (an indicator of potential item redundancy as it identified the variance predicted by the other items), and the Cronbach's alpha if the item was deleted (if the alpha increased when the item was removed then this is an indicator that it is harming the instrument's internal consistency). I noted that these reliability statistics did not add any new indicators of potential redundancy for any of the items. All corrected item-total correlations and squared multiple correlations were within acceptable limits and are presented in Table 4.14. No items, if removed, would increase the Cronbach's alpha. As a result, this additional information, although useful to check, did not make any real contribution to the process.

Therefore, as an alternative source of information to consider, I identified pairs of items with the highest correlations in the SI-20 inter-item correlation matrix. There were no inter-item correlations greater than .7, which would have indicated a high degree of correlation. Instead I made a note of all inter-item correlations greater than .6, identifying eleven pairs of items with a correspondence of at least 60%. These pairs are presented in Table 4.15. I used this information as a

confirmatory check of items that I had identified in my review of the data in Table 4.14 as potentially redundant.

Table 4.15. SI-20 inter-item correlations <.6

Item 1	Item 2	*r*
2. I have condemned myself for my attitudes or behaviour	3. I have tried to be what others think I should be	.62
4. I have trusted my own reactions to situations	20. I have felt it is all right to be the kind of person I am	.60
11. I have listened sensitively to myself	**16. I have felt true to myself**	.60
11. I have listened sensitively to myself	**17. I have been able to resolve conflicts within myself**	.68
13. I have lived fully in each new moment.	**16. I have felt true to myself**	.61
13. I have lived fully in each new moment.	**17. I have been able to resolve conflicts within myself**	.63
13. I have lived fully in each new moment.	20. I have felt it is all right to be the kind of person I am	.62
16. I have felt true to myself	**17. I have been able to resolve conflicts within myself**	.62
16. I have felt true to myself	20. I have felt it is all right to be the kind of person I am	.63
17. I have been able to resolve conflicts within myself	20. I have felt it is all right to be the kind of person I am	.66
19. I have accepted my feelings	20. I have felt it is all right to be the kind of person I am	.68

Note. **Bold** = items removed following analysis.

Following Bond & Fox's (2015) recommendations, I adopted an open, evaluative approach that allowed me to weigh up multiple factors in my decision-making process. My aim was to create a new shorter version of the SI that contained well-fitting items with evidence of relative stability (e.g. across versions, sub-samples, and pre-post functioning) and sensitivity to change (pre-post effect size) that represented a wide range of item difficulty.

I started by thinning out items with closely matched measures of item difficulty with the intention of using the data in Table 4.14, and if required, Table 4.15, to enable me to choose one item of the pair to remove from the instrument. I

identified four pairs of items that met this criterion: items 7 and 15; items 17 and 2; items 19 and 16; and items 8 and 6.

Item 7 and item 15. These two items had measures of .48 and .46, respectively. On inspection of Table 4.14, I noted that item 7 (*I have looked to others for approval or disapproval*) was an underfit to the model, suggesting that person responses on the item tended to be erratic. This finding was supported by the evidence of considerable disordered scoring identified when examining the SA sub-sample data. Its PMC (.52) was the lowest of the group, tending to support the view that it was not a strong member of the item group. This was supported by the low, but still viable, corrected item-total correlation (.47). The one factor in favour of keeping item 7 was the evidence of a large pre-post effect size. In comparison, item 15 (*I have hidden some elements of myself behind a 'mask*') had demonstrated a good fit to the model, a moderate PMC (.63) and no movement between strata or disordered scoring identified within the SA sub-sample data. Its main weakness was a low pre-post effect size. Neither item was present in Table 4.15 indicating no correlation greater than .6 with any other item. On balance, I decided to remove item 7 as it had more negative points (misfit, relatively low correlation with the instrument as a whole, item stability) than item 15 (effect size).

Item 17 and item 2. This pair of items clustered together with measures of .42 (item 17) and .41 (item 2). Table 4.14 indicated that item 17 (*I have been able to resolve conflicts within myself*) was overfitting and was found to have one of the highest PMCs (.76) of the items being assessed. This was mirrored by its corrected item-total correlation (.74) and squared mean correlation (.67). Movement between strata was identified within the SA sub-sample data. A positive result for this item was that it had been found to have a large pre-post effect size. However, this was also true for Item 2, the other item in the pair. In addition, item 2 (*I have condemned myself for my attitudes or behaviour*) was a good fit to the model, with a moderate PMC (.68) and no movement or disordered scoring noted in the SA sub-sample. Given these relative strengths, I made the decision to discard item 17. This

decision was supported by the presence of item 17 in four inter-item correlations presented in Table 4.15.

Item 19 and item 16. These two items fell on either side of the item measure mean: .01 and -.01, respectively. Item 19 (*I have accepted my feelings*) was one of the four items that I had returned to SI-20 having been removed from the Strathclyde Inventory when SI-16 was created. However, it did not fulfil my intended purpose: that it would have low item difficulty and therefore provide an additional item at the lower end of the scale. Instead, it was identified as an overfitting item, suggesting that it was not bringing anything new to the instrument. This was reflected in its moderately high PMC (.70). There was no information pertaining to potential movement or disordered scoring within the SA sub-sample as the item had not been included in SI-16 but I had been able to use SI-31 data to calculate a medium pre-post effect size. The second item in the pair, item 16 (*I have felt true to myself*) was also an overfitting item, with the highest PMC (.78) in the group. There was some movement between strata identified in the SA sub-sample data. The biggest concern in relation to this item was that I found it to have a negative effect size when tested with pre-therapy and post-therapy data. In weighing up the full range of this evidence gathered for these two items, I noted that their position on the hierarchy of item difficulty, as indicated by their measures, served to bridge two strata rather than separate them to create a space for development. As a result, I decided to remove both items: item 19 because it had not achieved its purpose on being returned to the instrument and item 16 because the evidence of a negative pre-post effect size indicated that it was a problematic item. This decision was also supported by the presence of both items within the inter-item correlations presented in Table 4.15

Item 8 and item 6. The final pair of items in this initial examination were situated at the bottom of strata 4, with measures of -.14 and -.15, respectively. Item 8 (*I have expressed myself in my own unique way*) was a good fit to the model, with a moderate PMC (.67). I had found it to have a medium pre-post effect size. However, my analysis of the SA sub-sample data had found that this item moved

between strata and had disordered scoring, providing evidence that it may lack stability. Item 6 (*I have felt afraid of my emotional reactions*) also had some weaknesses: it was an underfitting item with a relatively low PMC (.57) and a medium pre-post effect size. There was no evidence of movement or disordered scoring within the SA sub-sample data. I considered the content of each item. My sense was that the description 'my own unique way' included in item 8 could be off-putting and unnatural for some individuals. At the same time, I considered that the experience being described by 'afraid of my emotional reactions', the focus of item 6, might be sufficiently captured by item 12 (*I have felt myself doing things that were out of my control*) and item 4 (*I have trusted my own reactions to situations*). While both items had apparent problems, I made the decision to remove item 6 in the first instance, and to reserve item 8 as a candidate for possible removal at a later stage.

Item 5. Having completed my review of items with closely matching measures, I reviewed the statistics for individual items within strata 4 and 5, in the first instance, as I was reluctant to remove the lower difficulty items in strata 3. Item 5 (*I have experienced very satisfying personal relationships*) stood out as the one remaining underfitting item within these two strata. Table 4.14 indicated that it had a relatively low PMC (.59) and the lowest squared multiple correlation (.38) of the group. While this last finding might suggest low redundancy as it means that only 38% of the item's variance is predicted by the other items, it may also support the view that this item fits less well within the overall measure. Certainly, in my opinion, the content of the item is open to broad interpretation and is less closely related to, and indicative of, the theoretical construct as other remaining items. I had also found a small pre-post effect size for this item. Having considered this range of evidence, I decided to remove item 5.

Item 10. I was reluctant to remove items from the lower end of item difficulty, in particular those that I had re-introduced to SI-20 for this purpose, but had included strata 3 in this evaluation process as it contained two items: one that had been present since SI-16, item 10 (*I have made choices based on my own*

internal sense of what is right) and also item 12 (*I have felt myself doing things that were out of my control*), one of the four that I had re-introduced. The two items did not share closely matched measures (-.90 and -.83, respectively) so I did not consider them to be a pair. However, I had held doubts about the theoretical fit and/or interpretation of item 10 since my original analysis of the SI-16. Even with the addition of items with content that better fit my understanding of the experience of low congruence, this item still seemed out of place to me. In addition it had a relatively low PMC (.54), a small pre-post effect size, and I had found evidence through carrying out, first, a differential test functioning analysis then a differential item functioning analysis, that the functioning or meaning of this item changed, becoming more difficult, for participants at post-therapy (see chapter 5). Having taken all of these points into consideration, I decided that I would reserve item 10 as a candidate for possible removal at a later stage. My main concern in removing item 10 was that this may increase the distance between strata 3 and 4.

Testing alternative brief versions. Developing measures using Rasch measurement is an iterative process of creating and comparing alternative possibilities. Yan, writing in Bond & Fox (2015, pp.189-191), suggested the use of seven indicators to assess and select the best fitting version: practical considerations, item infit and outfit mean squares, point measure correlations, Rasch person and item reliability, variance explained by the measure, category function, and influence of underfitting persons. Therefore, using the results of my review of items, I created three alternative brief versions and produced their fit statistics for comparison. The three alternative versions were:

- SI-14: a 14-item version. Items removed: 5, 6, 7, 16, 17, 19.
- SI-13: a 13-item version. Items removed: 5, 6, 7, 8, 16, 17, 19.
- SI-12: a 12-item version. Items removed: 5, 6, 7, 8, 10, 16, 17, 19.

I have presented a comparison of the fit statistics and variance explained by the measure for SI-20 and the three alternative brief versions in Table 4.16 and a comparison of the item misfits across versions in Table 4.17.

Table 4.16. Comparison of fit statistics for SI-20 and three alternative brief versions: SI-14, SI-13 and SI-12

	SI-20	SI-14	SI-13	SI-12
Person reliability	.91	.87	.86	.85
Person separation (*G*)	3.18	2.64	2.50	2.42
Person strata*	3.91	3.19	3.00	2.89
Item reliability	.98	.98	.98	.98
Item separation (*G*)	6.62	7.68	7.94	7.83
Item strata*	8.49	9.91	10.25	10.11
Variance explained by the measure (%)	61.1	64.1	64.1	63.0

Note. * = ((4*G* – 1)/3).

As Table 4.16 illustrates, there was a steady decrease in the person reliability index, separation and strata as the number of items reduced. The lowest (SI-12; person reliability = .85, person separation = 2.42, person strata = 2.89) was still well within the acceptable limits as an indicator of good fit to the Rasch model and capable of identifying approximately three distinct strata of person ability. The item reliability index remained constant at .98, providing reassurance that the same group of items were likely to perform in the same way with another sample of participants of similar ability (Bond & Fox, 2015, p.363). Item separation and strata increased as the number of items decreased, reflecting the increasing differentiation in difficulty between the items. Finally, it can be seen that the variance explained by the measure improved as the instrument decreased in size, with marginally more variance explained by SI-14 and SI-13 (both 64.1%) than by SI-12 (63%).

Table 4.17 presents the infit mean square (IMSq), outfit mean square (OMSq) and point mean correlation (PMC) for each item across each version. This shows a similar picture. As the number of items reduced, version upon version, the mean squares of the more overfitting items increased and underfitting items decreased so that by SI-13 all items were well within the recommended range (.6 – 1.4). This trend is reflected by the PMCs: the correlation of underfitting items to the instrument, as a whole, increased slightly (<.04) and decreased slightly (<.03) for overfitting items.

Table 4.17. Comparison of item misfit statistics for SI-20 with three alternative brief versions: SI-14, SI-13 and SI-12

Item	SI-20			SI-14			SI-13			SI-12		
	IMSq	OMSq	PMC	IMSq	OMSq	PMC	IMSq	OMSq	PMC	IMSq	OMSq	PMC
12	1.45	1.41	.52	1.49	1.46	.52	1.40	1.37	.54	1.39	1.35	.56
7	1.39	1.49	.52	-	-	-	-	-	-	-	-	-
5	1.38	1.36	.59	-	-	-	-	-	-	-	-	-
6	1.37	1.33	.57	-	-	-	-	-	-	-	-	-
18	1.29	1.21	.53	1.32	1.23	.53	1.27	1.18	.54	1.26	1.17	.56
10	1.14	1.20	.54	1.13	1.16	.55	1.14	1.17	.55	-	-	-
14	1.14	1.16	.42	1.12	1.14	.44	1.13	1.16	.44	1.17	1.21	.44
15	1.13	1.12	.63	1.11	1.09	.65	1.08	1.07	.66	1.09	1.07	.65
1	1.08	1.12	.55	1.03	1.08	.58	1.04	1.08	.57	1.06	1.09	.57
9	1.02	1.01	.64	1.02	1.01	.65	.99	.99	.66	.97	.96	.66
3	.97	.96	.68	.96	.95	.69	.93	.92	.70	.92	.91	.70
2	.90	.93	.68	.90	.93	.69	.87	.90	.70	.87	.90	.70
8	.84	.83	.67	.81	.80	.70	-	-	-	-	-	-
11	.81	.84	.68	.82	.86	.69	.83	.86	.68	.87	.91	.66
19	.79	.78	.70	-	-	-	-	-	-	-	-	-
20	.79	.80	.75	.83	.84	.74	.84	.85	.74	.88	.90	.72
13	.73	.73	.72	.78	.79	.71	.79	.80	.70	.78	.80	.69
4	.67	.66	.71	.69	.68	.71	.68	.68	.71	.72	.72	.69
17	.59	.60	.76	-	-	-	-	-	-	-	-	-
16	.54	.55	.78	-	-	-	-	-	-	-	-	-

Notes. IMSq = infit mean square. OMSq = outfit mean square. PMC = point mean correlation.

Only item 14 (*I have been aware of my feelings*; strata 2) countered the trend by becoming slightly less well fitting in SI-12: SI-20 (IMSq = 1.14; OMSq = 1.16); SI-14 (IMSq = 1.12; OMSq = 1.14); SI-13 (IMSq = 1.13; OMSq = 1.16); SI-12 (IMSq = 1.17; OMSq = 1.21). In contrast, its PMC (.42; the smallest within the group) rises to .44, indicating that its relationship to the instrument as whole is marginally increasing.

These statistics showed that it was possible to reduce the items contained in the SI and not harm (indeed, slightly improve) its overall fit to the Rasch model. For practical reasons, the SI-12 appeared to be the most attractive version to adopt, because it was the briefest. However, there were two final tests that I wanted to check: first, that the items are sufficiently distributed across the instrument to fit persons across the range of ability; and second, that the content of the remaining items continued to make sense theoretically, without any obvious loss of meaning. Both of these issues can be considered by looking once more at an expected score item-person matrix, this time for the proposed new brief version of the Strathclyde Inventory, SI-12 (Figure 4.7).

As explained earlier in this chapter, the numbers underneath the matrix represent the participants in the sample according to their mean logit measure on the instrument. The numbers are presented vertically so, for example, the 1 and 2 above the M indicates that mean measures for 12 participants matched the mean for persons (.07). There are two extreme outliers; one at the high, one at the low end of the matrix. In addition, there are two participants whose mean measure was more than two standard deviations lower than the mean ('T' on the left) and six participants whose mean measure was more than two standard deviations higher than the mean ('T' on the right). Therefore, in a sample of 216 participants, ten (4.6%) have measures on the instrument that are too high or too low to be well represented.

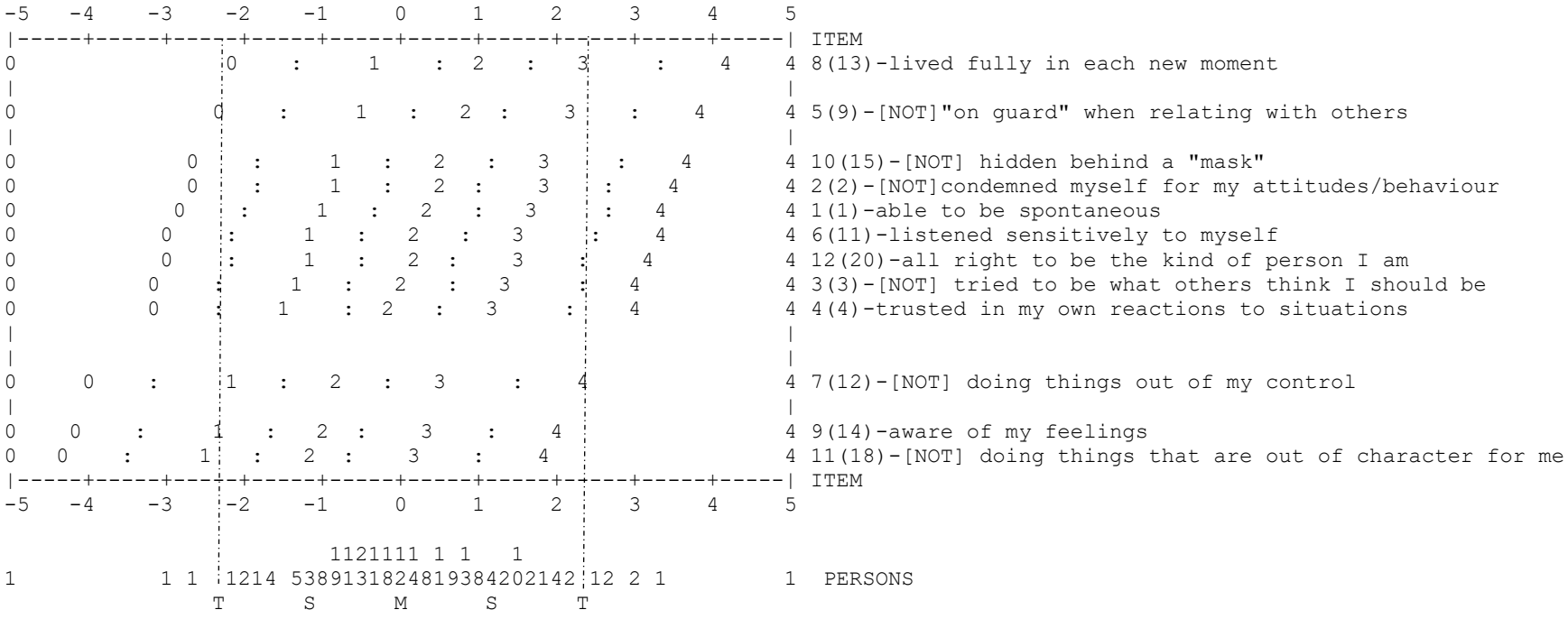


Figure 4.7. SI-12: Expected score item-person matrix

Notes. As for Figure 4.3. Dot/dash lines highlight the boundaries of two standard deviations above and below the mean.

The measures of the remaining participants fit comfortably within the boundaries of two standard deviations above and below the mean (highlighted by the two vertical dot-dash lines on Figure 4.7), indicating that they are well represented by the items comprising the SI-12.

Finally, I checked the content of items in the SI-12 hierarchy to assess its fit with the theoretical construct it was intended to measure. I found that the simplest way to approach this was to reflect on the relevance of the hierarchy as a description of a growing capacity for congruent functioning. Thus, the hierarchy of items presented the proposition that an individual may first experience a change in **self-awareness**, shifting from a sense that they are doing things that seem to be *out of character* to greater understanding and ownership of their behaviour. Part of this process is likely to involve becoming more *aware of their feelings* and as a result feeling *less out of control*. This experience of increasing self-awareness may develop into one of greater **self-trust**: for example, *trusting in their own reactions* to situations and feeling less need to *be the way others think they should be*. Continuing along the pathway, this increasing self-trust leads to increasing **self-acceptance**: feeling that *it's all right to be the kind of person that they are* and as a result become more able and willing to *listen sensitively* to their own needs. The pathway suggested that increased self-acceptance leads on to a growing sense of freedom, more **openness to self** and, gradually**, openness to others**: an individual may notice that they are able to be *spontaneous*, less *condemning of their attitudes and behaviour*, and having less *need to hide some elements of themselves behind a mask*. In due course, an individual may notice that they no longer felt the need to be *'on guard' when relating to others*, that they can let go of their former fears and *live fully in each new moment*.

I considered this to be a credible description of the process of developing a greater capacity for congruent functioning, consistent with my understanding of the theory and also my experience in therapeutic practice. As a result, I made the decision to accept SI-12 as the final product of this first study. A copy of the SI-12 has been included at Appendix E.

Chapter Summary

This chapter presented my investigation of the internal structure and precision/reliability of the SI when using data collected from a UK-based clinical population. I found evidence of excellent internal consistency whether analysing scores obtained from the whole sample, two subsamples (a general clinical population and participants meeting the diagnostic criteria for social anxiety), and two versions (SI-31 and SI-16), clearly replicating the results of previous studies of the SI using data collected from clinical and non-clinical populations (Freire, 2007; Zech et al., 2018). This demonstrates that scores obtained on the SI are measuring a coherent construct from the perspective of classic test theory.

However, going beyond previous studies, I used Rasch measurement to investigate the validity and precision/reliability of the SI more broadly and more specifically, assessing the five-category rating scale, the model fit and dimensionality of the instrument. I will highlight three key findings from this study.

First, I demonstrated that the 5-category rating scale worked well. As an instrument intended for use in clinical settings, it is essential that participants can distinguish categories within the rating scale and select the category that best fits their personal experience. Although 5-category rating scales can often cause difficulties for participants (e.g., Elliott et al., 2006), this was not the case for the participants who completed the SI-16 or its earlier 31-item version, within this study.

Second, I confirmed that the group of reverse scored items does not represent a separate factor or dimension within the instrument and that the SI (whether the SI-16, SI-20 or SI-12 version) represents a unidimensional construct acceptable to the stringent Rasch model. This finding provides a definitive answer to a long-standing question (Freire, 2007; Zech et al, 2018), and supports Rogers' proposition that the process of change in psychotherapy exists on a continuum "from fixity to changingness, from rigid structure to flow, from stasis to process" (1961/1967, p.131).

Third, in using Rasch analysis to investigate the difficulty of items contained in the SI in far greater detail than has been done before, I discovered that it was possible to distinguish meaningful levels, or degrees, of congruent functioning. Indeed, more than that, perhaps the most interesting discovery made in this study is the potential identification of a developmental pathway captured by the SI. Through evaluating the relative difficulty participants had in responding to different items on the SI, the application of Rasch measurement enabled me to tentatively map a process of developing congruent functioning: self-awareness, self-trust, self-acceptance, openness to self and others. This finding is generally consistent with Rogers' process conception of psychotherapy (1961/1967, pp.125-159) but goes further by highlighting the relational nature involved in the development of congruent functioning in which the struggle to trust and be open with others remained a concern, to some degree, for the majority of participants. This finding highlights a potential implication for those therapists who may assume that client trust is something that can be easily won before the 'real work' begins. The relevance of this potential pathway for the development of congruent functioning is tested further in my third study, reported in Chapter 6.

In addition to establishing evidence of the validity and reliability/precision of scores collected with the Strathclyde Inventory from a UK-based clinical population, it became apparent during the study that I could also develop the instrument to make it more user-friendly for test-takers. Having increased the number of items in the instrument in order to improve its validity, I carried out a series of analyses to reduce the number of items, without losing validity and reliability/precision. As a result, I produced a 12-item version of the Strathclyde Inventory; a brief outcome measure of congruent functioning, robustly validated for use with heterogenous UK-based clinical populations and ready for testing further afield.

Chapter 5: An Investigation of Change in Scores on the Strathclyde Inventory over the course of Therapy with Data Collected from a UK-based Clinical Population

Chapter Overview

In the previous chapter I investigated the evidence that supports the interpretation of a score collected on the Strathclyde Inventory (SI) as a measure of congruent functioning when completing the instrument on a single occasion. For my next study, presented in this chapter, I conducted a series of analyses to investigate change in scores on the SI collected from participants on two or more occasions over the course of therapy. Confirming that the SI is sensitive to change in scores over the course of therapy, including the degree to which change in scores on the SI converges or otherwise with other instruments, is an important aspect of its validation for use as an outcome measure with a clinical population. Ogles (2013, p.141) has noted that examination of outcome measures for sensitivity to change is a "growing trend" within the literature. Therefore, this study was designed to make a significant contribution towards answering my first overarching research question: are scores on the Strathclyde Inventory a valid measure of congruent functioning when used with a UK-based clinical population? At the same time, it was my intention to design this study in a way that would enable me to use my findings to discuss and extend theoretical understanding of congruent functioning and how this changes during therapy in order to address my second overarching question: do scores on the Strathclyde Inventory change over the course of therapy in a way that is consistent with person-centred theory?

I decided to examine the SI's sensitivity to change from a range of perspectives and designed my research questions with this in mind. I was interested in discovering what various methods for measuring change would highlight about the process of change in congruent functioning. I included conventional approaches to measuring change used in psychotherapy research (i.e. pre-post change in scores assessed for statistical significance, effect size and clinical significance) as well as less typical, more theoretically-driven analyses (i.e. shape of change across data

collection points, both group and individual, and change at individual item level). As a result, there are two main dimensions within this study: first, testing change on the SI calculated at group level and also at individual level; and, second, investigating change calculated solely on the basis of pre-therapy and post-therapy scores and change in scores calculated between data collection points *during* the process of therapy. Table 5.1 presents the relationship between my study-specific research questions, listed in full below, and these two dimensions (group-individual, pre/post-all) within the study.

Table 5.1. Relationship between study-specific research questions and two dimensions within approach to study

	Group Level	Individual Level
Pre-Post Data	2 (a, b, c, d) 4a 5 (a, b)	3 4a (i, ii)
All Data	4b (i, ii)	4b (ii, iii)

Research Questions

This second study was guided by the following questions:

1. Is there evidence that SI scores are temporally consistent prior to the start of therapy?
2. Do scores on the SI change over the course of therapy?
    a. Is the change in scores statistically significant?
    b. On average, what size is the change?
    c. How does change in scores on the SI compare with change in scores on the CORE-OM and PQ?
    d. Is there a relationship between change on the SI and (a) gender; (b) age; (c) number of sessions; (d) CORE-OM (non-risk and risk scales) pre-therapy scores; and CORE-OM (non-risk and risk scales) post-therapy scores?
3. According to their SI scores, what percentage of participants experience: (a) reliable change; (b) status change (i.e. clinical to nonclinical); and (c) clinically significant change (as defined by Jacobson & Truax, 1991)?

4. What is the shape of change in SI scores over the course of therapy (i.e. linear, nonlinear)?
    a. How well does pre-post change in SI scores fit a linear model?
        i. Do the scores of some participants exert an undue influence over the linear model?
        ii. Which participants had scores that changed more or less than expected based on their pre-therapy scores according to a linear model?
    b. What is the general pattern of change over time?
        i. Does change in SI scores over the course of therapy fit a nonlinear model?
        ii. What proportion of participants have scores that reliably increased or decreased between data collection points during therapy?
        iii. What happened to participants who continued in therapy after a data collection time point at which SI score suggested reliable deterioration?
5. In what ways do participants' scores on individual SI items change over the course of therapy?
    a. Are some items within the SI more sensitive to change?
    b. Do any SI items change in meaning for participants between pre- and post-therapy?

Method

Study Design

In this study, I returned to the same archived dataset that formed the basis for my first study but made the decision to include only the data collected from clients who participated in the PB protocol. Diverse in their reasons for accessing therapy, it was likely that this sample would reflect a typical heterogenous client population accessing person-centred therapy in routine practice. The data collected from them was done so as part of a single protocol organised on the principles of an

open clinical trial (Barker et al., 2016, p.147), also known as a *one group pretest-posttest design* or *within-subjects* study (Cook & Campbell, 1979, p.99), with *repeated measures across treatment* (Comer & Kendall, 2013, p.27) and provided data on the SI and two other outcome instruments, collected at a range of time points from a relatively large and varied sample of counselling clients. As a result, I excluded the data collected from clients in the SA protocol, a shorter and specialised protocol (e.g. diagnostic screening, additional outcome instruments, mid-therapy data collected after the eighth session, shorter number of sessions), which may have introduced additional, unintended variables into my investigation.

Participants

Clients. Participants were clients of the Strathclyde Counselling and Psychotherapy Research Clinic ('the research clinic') who accessed person-centred therapy between 2007 and 2017 within the generalist 'Practice-Based' (PB) protocol. The dataset for Study 2 contained all participants in Study 1 who completed the SI on at least two occasions during their therapy, including any additional observations collected since I extracted the data for Study 1, and any new participants who had contributed at least two observations by the time I commenced data analysis for Study 2. If a participant had left therapy without taking part in an end of therapy interview, resulting in them not completing the SI at the end of therapy, or were still in therapy, I followed a *last observation carried forward* approach (Comer & Kendall, 2013, p.34) in which their last score on the SI was treated as their post-therapy score.

The majority of participants were female (144, 64%; male = 80, 35.6%; non-binary = 1, .4%) and white British/European (N=215, 95.5%; Asian Indian/Pakistani = 4, 1.8%; Other = 6, 2.7%) with an age range of 18-67 years (M = 35.53, SD = 11.7).

Participants in the PB protocol at the research clinic during this period were offered up to 40 sessions of therapy. This maximum could be exceeded if the client and therapist agreed that it would be useful to continue the therapy and if the client was willing to pay a low-cost sessional fee. The participants in the Study 2 dataset attended between 3 and 70 sessions of counselling (mean = 25.21; median

= 23; mode = 40). All participants consented to take part in research activities alongside the counselling process.

Researchers. The data were collected by students and staff who participated as researchers in the Research Clinic between 2007 and 2017. I contributed to the data collection process for the PB protocol as Research Clinic coordinator from 2013 onwards. Professor Robert Elliott was the principal investigator.

Data Collection

Strathclyde Inventory. As outlined in Chapter 4, participants in the PB protocol completed the SI at regular data collection points during the therapeutic process: before therapy began, following every 10th session of therapy, and at the end of therapy. The number of data collection points at which the SI was completed per participant depended on the number of counselling sessions in which they participated. There were 776 SI observations included in the Study 2 dataset collected at nine different data collection points (see Table 5.2).

Table 5.2. Study 2 dataset presented by data collection point

Data collection point	Type of analyses	
	Pre-Post (N)	Shape of Change (N)
Pre-therapy	225	225
1st session	44	-
Mid-1	132	199
Mid-2	82	141
Mid-3	54	91
Mid-4	9	58
Mid-5	4	-
Mid-6	1	-
Post-therapy	225	-
Total	776	714

The range of observations collected per participant was 2 - 8 (mean = 3.45; median = 3; mode = 2). As described earlier, where participants had not completed a post-therapy SI score, I recoded their latest observation as 'post-therapy' for analyses in which I was comparing pre-therapy and post-therapy scores. This enabled the dataset in these analyses to have an equal number of pre-therapy and post-therapy

observations. For later analyses in which I investigated non-linear shape of change, I included scores at mid-points across therapy up to 40 sessions, using the original data collection point codes and, in addition, inserting post-therapy data at the closest mid-therapy data collection point in order to assess change in relation to time up to the post-40 sessions point, which was the usual maximum number of sessions offered in the research clinic.

The instrument was administered to participants by their researcher at each data collection point along with CORE-OM (Evans et al., 2002), designed to measure general concepts of distress and functioning, and the Personal Questionnaire (PQ; Elliott et al., 2016), an individualised instrument created by the client at the intake interview to identify specific difficulties that they wished to address in therapy. These instruments formed the standard set of outcome measures for the PB protocol. The SI was presented to participants within this battery of outcome measures and was usually the second or third of the three instruments completed on each occasion.

Three different versions of the SI were included in this dataset: the 31-item version (SI-31; Elliott & Rodgers, 2007), which was used from 2007-2012; the 16-item version (SI-16; Elliott & Rodgers, 2012c), which was used from 2012-2016; and the 20-item version (see Chapter 4). Ninety participants provided data using SI-31 only; 90 participants worked only with SI-16; and 3 participants contributed observations using SI-20 only. Forty two participants made a transition between versions during their therapy process: either SI-31 to SI-16 (N=12) or SI-16 to SI-20 (N=30).

Participants were asked to read each statement on the SI, consider how often it had been true for them during the last month, then mark the box that was closest to this on the 5-point rating scale.

They completed two other outcome measures, the CORE-OM and the PQ, at the same data collection points.

CORE-OM. The Clinical Outcomes in Routine Evaluation – Outcome Measure (CORE-OM) was designed to be an "acceptable, standardized outcome measure [that could be used] in a wide range of practice settings" (Evans et al., 2002, p.51). The CORE-OM is a 34-item self-report instrument, which asks test-takers to read each statement and report how often they have felt that way in the last week using a 5-category Likert scale: not at all, only occasionally, sometimes, often, most or all of the time (CORE System Trust, n.d.). The items were selected to represent four domains: subjective well-being (four items; e.g. item 4 – *I have felt OK about myself*), problems/symptoms (twelve items; e.g. item 11 – *tension and anxiety have prevented me from doing important things*), life functioning (12 items; e.g. item 19 – *I have felt warmth or affection for someone*), and risk to self and others (self, four items; others, two items: e.g. item 34 – *I have hurt myself physically or taken dangerous risks with my health*; item 22 – *I have threatened or intimidated another person*).

The internal consistency of the CORE-OM is excellent (e.g. Evans et al., 2002, α=.94; Connell et al., 2007, α=.91) as is its test-retest reliability after one week (Evans et al., 2002, Spearman's ρ=.90). The recommended clinical significance cut-off score is 1.0 and the RCI minimum value for a clinical population is .59 ($p < .05$) (Connell et al., 2007).

While acknowledging that the distinction between four domains has been found useful by managers and clinicians, Lyne, Barrett, Evans & Barkham (2006) reported that the CORE-OM is best scored as two scales: one, psychological distress (α=.93), containing the three domains of subjective wellbeing, problems/ symptoms, and life functioning (28 items), and the other, risk (α=.77; 6 items). They noted "relatively little differentiation" between the three non-risk domains (Lyne et al., 2006, p.200) and proposed three reasons for the apparent difficulty in distinguishing between problematic symptoms and functioning in typical counselling populations that this indicated: difficulties in functioning is often evidence of psychological distress; people accessing counselling in outpatient services tend to be functioning at a manageable level compared to people with

severe and enduring psychological difficulties, therefore more sensitive measures may be required; and other, more successful, measures of functioning include indicators that refer to actual behaviours rather than accompanying distress.

As originally conceived, CORE-OM has been widely adopted in counselling settings permitting large-scale research into the effectiveness of counselling in routine practice within the UK (e.g. Barkham et al., 2001; Stiles et al., 2006; Stiles et al., 2008).

The internal consistency of scores on the CORE-OM and its two scales in the dataset used in this study was reasonably consistent with the populations in previous studies: slightly higher for CORE-OM (34 items; α=.95) and CORE-OM non-risk scale (28 items; α=.95) and somewhat lower for CORE-OM risk scale (6 items; α=.69).

Personal Questionnaire. The simplified Personal Questionnaire (PQ; Elliott et al., 2016) is a client-generated outcome measure (CGOM), representing an idiographic approach to outcome measurement. In 2014, Sales & Alves (cited in Elliott et al., 2016) reported that the PQ was the most popular CGOM, used in eleven published studies. The PQ is an individualized instrument created by the client before beginning therapy in which they create a list of specific difficulties that they wish to address in therapy. They rate how much each problem has bothered them during the past seven days using a seven-category rating scale: not at all, very little, little, moderately, considerably, very considerably, maximum possible.

Having analysed data collected from the PQs of 455 participants in studies conducted in the USA, Portugal and Scotland, including participants in this dataset, Elliott et al. (2016, pp.274-275) reported: normative characteristics of PQ scores (approximately 10 items, mean pre-therapy scores (i.e. initial severity) = 5, equating to 'considerably' on the 7-category rating scale); good internal consistency both between-clients and within-client (in the range α = .70 - .80); consistent temporal reliability (*r* = .57); strong correlations with a selection of standardised outcome measures (typically in the range *r* = .30 - .60); large pre-post standardised mean

differences in scores (d = 0.8 – 1.7). They recommended a clinical significance cut-off score of 3.25 and an RCI minimum value of 1.5 ($p < .05$).

Data Analysis

The research questions listed at the beginning of this chapter guided the analyses that I conducted. I have described the specific procedures that I followed in the next section alongside the results. I used SPSS software (version 23) to conduct the statistical analyses and Winsteps (version 3.62.1; Linacre, 2006) for the Rasch analyses.

Results

Is there Evidence that SI Scores are Temporally Consistent Prior to the Start of Therapy?

A reliable outcome measure must demonstrate that it is *sensitive to change* but there must also be evidence that it can produce stable and consistent scores when completed by the same individual on more than one occasion when *no change* in the construct is likely to have occurred. In the case of the SI, it would be expected that there would be little, if any, change in a client's congruence in the immediate period prior to beginning therapy.

In the dataset collected for Study 2, 44 clients had completed the SI on two occasions before commencing therapy: the first at their pre-therapy intake interview, the second immediately before their first counselling session. The median duration between the two time points was 15 days (mean = 44 days; range = 2 – 144 days). I assessed test-retest reliability by carrying out a Pearson's correlation between the scores at the two data collection points (Barker et al., 2016, p.61). This calculation found a large correlation between scores collected at the two time points: r = .81, $p < .01$. This result provides evidence that the scores on the SI were temporally consistent for these clients prior to commencing therapy.

Do Scores on the SI Change over the course of Therapy?

In this section, I present the results of my analyses of change in aggregated SI scores over the course of therapy (i.e. at group level), first using the concepts of

statistical significance, to test the unlikeliness that change can be explained by measurement error, and effect size, to assess mean difference converted to a standardised metric. Next, I present evidence of the convergent validity of the SI, the PQ and the CORE-OM, using data collected from this dataset, by comparing effect sizes and pre-post correlations and conducting a two-tailed Pearson correlation to investigate the relationship between the pre-post difference in scores recorded for each of the three measures. Last, I present the results of two-tailed Pearson correlations to test if change on the SI varies according to gender, age, number of sessions or CORE-OM pre-therapy or post-therapy scores.

Is the change in scores statistically significant? The t-test is the standard analysis used to compare two means in order to identify if a statistically significant change can be detected. In this case, as this was a repeated-measures design, I chose to use the paired-samples t-test (Field, 2013, p.378).

This calculation showed that, on average, the 225 participants' mean SI score at pre-therapy was 1.79 (SD = .65, SE = .04). At the end of therapy their aggregated mean SI score was 2.47 (SD = .78, SE = .05). The difference, -.68, was significant $t(224) = -13.53$, $p < .001$. The correlation between pre-therapy and post-therapy scores was .46.

In order to examine sensitivity to change across SI versions, I conducted paired-samples t-tests using the data collected on SI-31 and SI-16 from participants who did not transition across versions during their therapy. I did not analyse the data for SI-20 as the sample was too small (N=3). The results showed that the change in the aggregated pre-post mean scores recorded for each version was statistically significant: SI-31 $t(89) = -8.35$, $p < .001$, $r = .48$; SI-16 $t(89) = -9.68$, $p < .001$, $r = .44$. I did not test the difference between the two correlations because it was clearly so small that it would take an enormous sample for the difference between versions to be statistically significant. These results are displayed in Table 5.3.

On average, what size is the change? I used this data to calculate the effect size (Cohen's *d*), also commonly referred to as the standardized mean difference, which provides a measure of the strength of the relationship between the aggregated mean scores at the two data collection points that is unrelated to sample size (Barker et al., 2016, p.229). Cohen's *d* is calculated by dividing the difference between the means of the two groups by their pooled standard deviation. Based on the N, M and SD for each group (pre-therapy and post-therapy) reported above, I calculated a pooled SD of .72 resulting in a finding of *d* = -.95; *d* is negative because higher scores on the SI indicate improvement. Using the same calculation method, I found the effect size for SI-31 was -.91 and for SI-16, -1.08. Following Cohen (1992), these can be considered very large effects, with the data collected on SI-16 indicating the largest effect. However, this difference in effect size between the two versions was only .17, a small effect, and unlikely to be statistically significant. The results are displayed in Table 5.3.

Table 5.3. Comparison of statistical significance, effect size and correlation of pre-post change across SI versions

	Pre-therapy		Post-therapy		*t*	Effect size *d*	Pre/post *r*
	N	Mean (SD)	N	Mean (SD)			
SI (all)	225	1.79 (.65)	225	2.47 (.78)	-13.53**	-.95	.46**
SI-31	90	1.93 (.62)	90	2.58 (.80)	-8.35**	-.91	.48**
SI-16	90	1.63 (.67)	90	2.39 (.73)	-9.68**	-1.08	.44**

Notes. ** = p <.001

How does change in scores on the SI compare with change in scores on the CORE-OM and PQ? I repeated these analyses for the pre-therapy and post-therapy data collected on the CORE-OM (separating non-risk and risk scales as recommended by Lyne et al., 2006) and the PQ. Not all participants in the dataset had completed all three instruments on both occasions.

First, I used a paired-samples t-test with the scores collected from 221 participants on the CORE-OM (non-risk scale; CORE-NR) to calculate an aggregated mean score at pre-therapy of 2.08 (SD = .71, SE = .05) and at post-therapy of 1.34 (SD = .83, SE = .06). The difference (.73) was significant $t(220) = 13.73$, $p < .001$, $r = .48$, and represented a very large effect ($d = .96$), almost identical to the effect size and pre-post correlation for the SI. This finding is interesting because, as Elliott (2001, p.64) demonstrated, measures of symptoms of psychological distress, such as CORE-NR, tend to be associated with much larger pre-post effect sizes than measures of experiential functioning like the SI. In contrast, these results indicate that the two measures detected a very similar degree of change in participants, echoing the strong relationship between scores on the SI and CORE-OM found by both Freire (2007) and Tashiro (2011).

Next, I conducted a paired-samples t-test with data collected from 221 participants on the CORE-OM's risk scale (CORE-R). The aggregated mean score at pre-therapy was 0.38 (SD = .52, SE = .03) and at post-therapy of 0.20 (SD = .41, SE = .03). The difference, 0.18, was significant $t(220) = 5.85$, $p < .001$, $r = .51$ and represented a moderate (small-medium) effect ($d = .38$). It is likely that the effect size is relatively small because of a restriction in the range of scoring at the lower end of the rating scale, known as a *basement effect.* In other words, most participants are likely to have started and ended therapy with a low rating on the risk scale.

Following the same procedure with data collected from 209 participants who completed the PQ, I calculated an aggregated mean score at pre-therapy of 5.15 (SD = .81, SE = .06) and at post-therapy of 3.43 (SD = 1.33, SE = .09). The difference, 1.72, was significant, $t(208) = 17.91$, $p < .001$, $r = .24$, and represented an extremely large effect (d = 1.56). This result is also consistent with Elliott (2001, p.64), in which individualised outcome measures were identified as being associated with the largest effect sizes.

A summary of the results for all three measures are presented in Table 5.4. Please note that an increase in scores on the Strathclyde Inventory indicates

improvement whereas on the other two instruments scores are expected to decrease.

Table 5.4. Comparison of statistical significance, effect size and correlation of pre-post change on the SI, CORE-OM (non-risk and risk scales) and PQ

	Pre-therapy		Post-therapy		*t*	Effect size *d*	Pre/post *r*
	N	Mean (SD)	N	Mean (SD)			
SI	225	1.79 (.65)	225	2.47 (.78)	-13.53**	-.95	.46**
CORE-NR	221	2.08 (.71)	221	1.34 (.83)	13.73**	.96	.48**
CORE-R	221	0.38 (.52)	221	0.20 (.41)	5.85**	.38	.51**
PQ	220	5.16 (.81)	214	3.42 (1.32)	17.91**	1.56	.24**

Notes. ** = p < .001. ES >= 0.8 = large effect; ES >= 0.5 = medium effect; ES >= 0.2 = small effect (Cohen, 1992).

As Table 5.4 demonstrates, the pre-post change in scores on all four instruments/scales is statistically significant. For all but the CORE risk scale, the change represents a very large effect, particularly for the PQ. This is a common finding on this instrument (Elliott et al., 2016). The correlation between pre-therapy and post-therapy scores on each of these three measures/scales matches the findings of Tashiro (2011: SI = .45; PQ = .23; CORE-OM = .47). The similarity in effect size between SI and CORE-NR (in the opposite direction consistent with the direction of scoring on each instrument) provides further evidence of the apparent overlapping relationship between the two measures noted by Freire (2007).

Next, I conducted a two-tailed Pearson correlation to investigate the relationship between the pre-post difference in scores recorded for each of the three measures. Although the SI scored in a different direction (a higher score is better) than the other three measures/scales, pre-post change on all four measures/scales was calculated so that an improvement in scores was represented by a positive number, and a deterioration in scores was represented as a negative number. The results are presented in Table 5.5.

Table 5.5. Correlation of pre-post difference in scores on the SI, CORE-OM (non-risk and risk scales) and PQ

	SI	CORE-NR	CORE-R	PQ
SI	-	.76**	.27**	.55**
CORE-NR	.76**	-	.46**	.66**
CORE-R	.27**	.46**	-	.36**
PQ	.55**	.66**	.36**	-

Note. SI = Strathclyde Inventory; CORE-NR = CORE-OM non-risk scale (28 items); CORE-R = CORE-OM risk scale (6 items); PQ = Personal Questionnaire; ** = p < .01 (2-tailed)

These results confirmed the previous findings: that there was a strong association between pre-post difference in scores for the SI and the CORE-NR (*r* = .76), and a somewhat smaller, more moderate relationship between pre-post difference in scores recorded by the SI and PQ (*r* = .55). The correlation of the pre-post difference in scores for the CORE-NR and PQ fell between these two contrasting positions (*r* = .66). These results suggest that the SI is measuring change of a different nature to the client-identified difficulties captured by the PQ, while the CORE-NR appears to bridge this gap, consistent with Lyne et al.'s (2006) finding that the CORE-OM cannot distinguish between psychological symptoms and functioning. However, a greater distinction between the SI and CORE-NR is indicated when the difference in pre-post scores on each measure is correlated with the CORE-R scale, a measure of risk: with the CORE-NR, a measure of general psychological distress, the correlation is .46; with the PQ, a measure of individualised distress = .36; and the correlation with the SI, by far the smallest, is .27. This finding suggests that the amount of change in congruent functioning has only a small association with the amount of change in level of risk during therapy; in other words, that the SI is measuring something distinctively different.

Is there a relationship between change on the SI and (a) gender; (b) age; (c) number of sessions; (d) CORE-OM (non-risk and risk scales) pre-therapy scores; and CORE-OM (non-risk and risk scales) post-therapy scores? I conducted two-tailed Pearson's correlations to test if there was any association between pre-post change in SI scores (first as a raw score indicating how much the score had changed,

then as a standardised residual indicating the degree to which the change was more or less than expected by the participant's pre-therapy score based on a regression analysis using the whole dataset, see later in this chapter) and five potential predictor variables of outcome: gender, age, number of sessions, and CORE-OM (non-risk and risk scales) pre-therapy scores (as indicators of clients' level of general psychological distress and risk when commencing therapy). In addition, I tested the relationship between change in SI scores (first as a raw score, then as a standardised residual) as potential predictor variables of outcome as measured by CORE-OM (non-risk and risk scales) post-therapy scores (i.e. indicators of clients' level of general psychological distress and risk when ending therapy). The results of these two series of correlations are presented in Table 5.6:

Table 5.6. Correlations between pre-post change in SI scores and gender, age, number of sessions and CORE-OM (non-risk and risk scales) pre-therapy and post-therapy scores

	Pre-post change (raw score)	Standardised residual (*Z* score)
Gender	.10	.05
Age	.03	.11
Number of sessions	.03	.03
CORE-NR pre-score	.20**	-.08
CORE-R pre-score	-.00	-.14*
CORE-NR post-score	-.54**	-.74**
CORE-R post-score	-.25**	-.40**

Note. * = $p < .05$ (2-tailed); ** = $p < .01$ (2-tailed)

As Table 5.6 shows, I did not identify any statistically significant association between pre-post change in SI scores (whether raw or standardised) and gender, age or the number of therapy sessions accessed. However I did find six correlations that were statistically significant: (1) a small positive correlation (.20; $r^2 = .04$) between CORE-NR pre-therapy scores and the raw score version of the change in SI scores (mean difference); (2) a small negative correlation (-.14; $r^2 = .02$) between CORE-R pre-therapy scores and the standardised residual for SI pre-post change; (3) a large negative correlation (-.54; $r^2 = .29$) between CORE-NR post-therapy scores and the SI raw score difference; (4) a very large negative correlation (-.74; $r^2 = .55$)

between CORE-NR post-therapy scores and the SI standardised residual; (5) a small-medium negative correlation (-.25; r^2 = .06) between CORE-R post-therapy scores and the SI raw score difference; and (6) a medium negative correlation (-.40; r^2 = .16) between CORE-R post-therapy scores and the SI standardised residual.

This indicates that the more distressed participants were at pre-therapy, the slightly more likely it was that their SI scores would increase by a large amount by the end of therapy (explaining 4% of the variance) but if there was more risk for participants at pre-therapy, then it was slightly more likely that their SI scores would change less than expected (according to the standardised residual) by the end of therapy (explaining 2% of the variance). In contrast, the larger the change in SI scores (raw difference) or actual change compared to expected change (standardised residuals) between the beginning and end of therapy, the greater the likelihood that participants' scores at post-therapy would indicate lower distress (on the CORE-NR: 29% or 55% variance, respectively) and risk (on the CORE-R: 6% or 16% of the variance, respectively). These results suggest that participants who recorded the most change in their SI scores during therapy were somewhat more likely to experience minimal or no distress and risk at the end of therapy.

Summary of section. The results of these analyses using aggregated data collected from 225 participants who accessed person-centred counselling at the Strathclyde research clinic demonstrates that:

- the change in mean SI scores between pre-therapy and post-therapy was statistically significant ($p < .001$; $r = .46$) and represented a large effect size ($d = .95$); there was no substantial difference between the results for versions SI-31 and SI-16 when analysed separately.
- the size of change and pre-post correlation for the SI was almost identical to the results for the CORE-OM (non-risk scale), replicating previous studies comparing the SI with the CORE-OM as a whole, but surprising given the different type of change that the two instruments set out to measure. According to Elliott's (2001) meta-analysis, an instrument measuring functioning (e.g. the SI) would be expected to demonstrate a smaller effect

size that an instrument designed to measure symptoms of distress (e.g. the CORE-OM).

- in contrast, the size of change and pre-post correlation for the PQ was different to that of the SI, and consistent with findings for individualised measures of distress (Elliott, 2001).
- evidence of convergence between the SI and CORE-NR and divergence between the SI, PQ and CORE-R (risk scale) was further supported by the strong association measured between the pre-post change in scores for the SI and CORE-NR(r = .76), the weaker association between pre-post change in scores for the SI and PQ (r = .55) and the weak association between pre-post change in scores for the SI and CORE-R (r = .27).
- the association (r = .66) detected between pre-post difference in scores on the CORE-NR and PQ suggests that the CORE-NR, rather than replicating the SI, overlaps the construct of congruent functioning, as measured by the SI, and the individualised client distress measured by the PQ. However, when the relationship between pre-post change in scores on these three measures/scales are tested with those of the CORE-R, a greater distinction between the constructs measured by the CORE-NR (r =.46) and SI (r =.27) becomes evident.
- initial higher levels of distress may have a small effect (r^2 = .04) on the likelihood of change in SI scores by the end of therapy, whereas initial higher levels of risk may have a small effect (r^2 = .02) on the likelihood of a less than expected change in SI scores by the end of therapy.
- a greater amount of, or more than expected, pre-post change as measured by the SI may have a large effect on the likelihood of participants recording low post-therapy scores of distress (raw difference: r^2 = .29; standardised residuals: r^2 = .55) and a small-medium effect on likelihood of low post-therapy scores of risk (raw difference: r^2 = .06; standardised residuals: r^2 = .16), as measured by the non-risk and risk scales of the CORE-OM.

According to their SI Scores, What Percentage of Participants Experienced (a) Reliable Change; (b) Status Change; (c) Clinically Significant Change (as Defined by Jacobson & Truax, 1991)?

In Chapter 2, I introduced the concepts of clinical significance and reliable change (Jacobson & Truax, 1991). In this section I will present the results for reliable change, status change (moving from one 'population' to another, according to scores on the measure, i.e. clinical to non-clinical) and clinically significant change recorded by the participants in my study according to pre-post change in their SI scores.

Before beginning it was necessary to identify the appropriate metrics to use for these analyses. Therefore, I combined the data collected in my dataset with that of previous studies (see Chapter 2) to calculate a standardised clinical significance cut-off score and reliable change indices that could be used in my study, and by other researchers interested in using the SI as an outcome measure within their work.

Calculation of standardised clinical significance cut-off score and reliable change indices. Jacobson and Truax (1991, p.13) proposed that a clinical significance cut-off score can be calculated using data from a dysfunctional population by defining functioning scores as those falling (in the direction of functionality) at least two standard deviations beyond the mean of the scores collected from the dysfunctional population. As Table 5.3 shows, the aggregated pre-therapy SI mean for my dataset is 1.79, with a standard deviation of .65. Therefore, according to Jacobson and Truax criterion A, this indicated a clinical significance cut-off score of 3.09 (i.e. 1.79 + (2 x .65)), which is considerably higher, and therefore more stringent, than earlier calculations of a clinical significance cut-off score for the SI using criteria B and C with data collected in previous studies (see Chapter 2). This is to be expected: Jacobson and Truax (1991) noted that criterion A tends to produce a more conservative cut-off score, whereas criterion B, calculated using non-clinical data, tends to result in a more lenient result as demonstrated by Freire (2007). The preferred method, when both clinical and non-clinical data is

available, is to use Criterion C. Therefore, to produce a standardised score, I gathered together the data used by Freire (2007), Folkes-Skinner (2011), Zech et al. (2018) and this study. I calculated weighted means on the required metrics: SI mean and standard deviation and test-retest for non-clinical participants; pre-therapy SI mean and standard deviation and test-retest for clinical participants. The results of these processes are displayed in Table 5.7.

Table 5.7. Calculation of weighted means

Study	SI mean score				Test-retest		
	N	M	WM	SD	N	T-R	WT-R
Non-clinical samples							
Freire (2007)[a]	399	2.79	1113.21	.54	77	.66	50.82
Folkes-Skinner (2011)[b]	18	2.88	51.84	.51	-	-	-
Zech et al. (2018)[c]	104	2.63	273.52	.48	104	.73	75.92
	119	2.91	346.29	.44	119	.63	74.97
	61	2.73	166.53	.48	-	-	-
	36	2.83	101.88	.86	-	-	-
Total	737		2053.27		300		201.71
Weighted M			**2.79**				
Pooled SD				.57			
Clinical samples							
Zech et al. (2018) [c]	15	2.13	31.95	.48	9	.46	4.14
	57	2.33	132.81	.71	56	.63	35.28
	10	2.71	27.1	.42	10	.01	.10
Stephen (this study)[d]	225	1.79	402.75	.65	44	.81	35.64
Total	307		594.61		119		75.16
Weighted M			**1.94**				
Pooled SD				.65			
Whole sample							
Total	1044				419		276.87
Pooled SD				**.60**			
Weighted M: T-R							**.66**

Notes. [a] = Version 2; [b] = Version 4; [c]= Version 5; [d] = Versions 4,6&7. WM = weighted mean of mean; WT-R = weighted mean of test-retest score; Pooled SD = pooled standard deviation.

Following Elliott's (2004) recommendations, I used these weighted means, first, to calculate a clinical significance cut-off score according to criterion C ((2.79 + 1.94)/2 = 2.36) then, using pooled standard deviations for the whole sample (based on the pooled standard deviations calculated for the non-clinical and clinical populations) and weighted means of test-retest when available, to calculate RCI minimum value metrics according to the equation below: *s* is the weighted mean standard deviation for the non-clinical population (.60); r_{xx} is the weighted mean test-retest for the combined population (.66); and z takes the value corresponding to the level of *p* required (e.g., for $p < .05$, $z = 1.96$).

$$RCI_{min} = z\left(s\sqrt{2(1-r_{xx})}\right)$$

Stephen & Elliott (2011, pp.237-238) discussed the issue of *standard of proof* when assessing client change in counselling research by comparing statistical principles with legal concepts – i.e. *beyond all reasonable doubt* ($p < .05$) and *on the balance of probability* ($p < .5$) - and advocated the appropriateness of the intermediate standard used in the American legal system for high stakes civil cases such as child custody matters: *clear and convincing evidence* ($p < .2$). This reassessment of statistical conventions in which an arbitrary default of 95% is used to conclude if change has been proven is consistent with the perennial criticism of the use of statistical significance as a dichotomous device, most recently outlined by Amrhein, Greenland & McShane (2019). With these arguments in mind, I decided to produce RCI values for the two higher levels of probability so that I could compare and assess the different results produced. Using the data displayed in Table 5.7, I calculated minimum RCI values: .97 (for $p < .05$) and .64 (for $p < .2$). These standardised scores have been used in this study for the purpose of calculating reliable change, status change and clinically significant change.

Calculation of reliable change, status change and clinically significant change for participants in this study. Having determined a clinically significant cut-off score and two RCI minimum values, I used these indices to identify for each

individual participant if their scores had increased (or decreased) between pre-therapy and post-therapy: first by at least .97, indicating that reliable change had occurred at the standard of $p < .05$; then by at least .64, indicating that reliable change had occurred at the standard of $p < .2$. I also checked which population range their scores matched at pre-therapy and then at post-therapy (clinical < 2.36 < non-clinical). Once I had completed these processes, I identified which participants had scores that indicated they had achieved reliable change and had crossed from the clinical range to the non-clinical range by the end of therapy: i.e. clinically significant change. Then I organised the participants' data into groups according to their classification for reliable change: i.e. reliable change (improvement), reliable change (deterioration) and no reliable change. Table 5.8 presents the results of these processes according to reliable change at the higher standard ($p < .05$).

Table 5.8. Reliable change ($p<.05$), status change and clinically significant change

	Reliable change (improvement)		Reliable change (deterioration)		No reliable change	
	N	%	N	%	N	%
Total	70	31.1	2	0.9	153	68.0
Status change (N=93; 41.3%)						
-Clinical to non-clinical	59	26.2*	-	-	29	12.9
-Non-clinical to clinical	-	-	2	0.9	3	1.3
No status change (N=132; 58.7%)						
-Clinical	6	2.7	-	-	85	37.8
-Non-clinical	5	2.2	-	-	36	16.0

Notes. RCI = .97($p<.05$); clinical cut-off point = 2.36; * = clinically significant change.

As can be seen, the scores of 31.1% participants indicated that reliable change (improvement) had taken place by the end of therapy. According to their scores, 59 of these participants (26.2% of the total number of participants) also moved from the clinical range at pre-therapy to the non-clinical range at post-therapy, and therefore can be identified as achieving clinically significant change over the course of therapy. Of the remaining participants whose scores demonstrated reliable improvement, 5 (2.2%) had scores in the non-clinical range at pre-therapy and therefore were not able to register clinically significant change, while the increase in

score for the other 6 participants (2.7%) was not enough to lift them from the clinical range by post-therapy.

The scores of two participants (0.9%) indicated reliable deterioration had occurred over the course of therapy, with both participants' scores crossing the threshold from non-clinical to clinical by the end of therapy.

According to my calculations, the largest group of participants (N=153; 68.0%) were those whose scores on the SI between pre-therapy and post-therapy did not change sufficiently to meet the criterion for reliable change ($p < .05$). The scores of 32 of this group (14.2% of all participants) were clustered around the cut-off between the clinical and non-clinical range and appeared to cross the threshold over the course of therapy. The scores of 85 of the remaining participants (37.8%) were firmly located in the clinical range, while the scores of another 36 participants (16.0%) started and ended therapy in the non-clinical range.

I was curious to know how these percentages might change if the standard for reliable change was adjusted to the $p < .2$ level so conducted the same calculations again, this time applying .64 as the RCI. Table 5.9 presents the results.

Table 5.9. Reliable change ($p<.2$), status change and clinically significant change

	Reliable change (improvement)		Reliable change (deterioration)		No reliable change	
	N	%	N	%	N	%
Total	105	46.6	7	3.1	113	50.3
Status change (N=93; 41.3%)						
-Clinical to non-clinical	75	33.3*	-	-	13	5.8
-Non-clinical to clinical	-	-	3	1.3	2	0.9
No status change (N=132; (58.7%)						
-Clinical	18	8.0	2	0.9	71	31.6
-Non-clinical	12	5.3	2	0.9	27	12.0

Notes. RCI = .64 ($p<.2$); clinical cut-off point = 2.36; * = clinically significant change.

As Table 5.9 shows, this change in RCI affected the classification of 40 participants whose change in scores had previously been designated as not indicating reliable change. Of these, 35 participants are now in the reliable change (improvement) group; the remaining five have moved into the reliable change (deterioration)

group. Sixteen of the participants who are now deemed to have experienced reliable improvement in scores can also be classified as having achieved clinically significant change, increasing the overall percentage of participants whose scores indicate clinically significant change to 33.3%.

The revised group of participants whose scores indicate reliable deterioration are now spread between those whose status changed from non-clinical to clinical (N = 3; 1.3%) and those whose status remained the same: clinical (N = 2; 0.9%) and non-clinical (N = 2; 0.9%).

The lowering of the standard for reliable change did not affect the classification of 113 participants (50.3%), approximately half of the participants in this study.

One final result that I noted was that 179 participants (79.6%) had pre-therapy scores in the clinical range, while the scores of the 46 remaining participants (20.4%) indicated that they were in the non-clinical range before commencing therapy. This replicates the percentage of non-clinical pre-therapy scores noted by Tashiro (2011; 20.5%) and, at almost one fifth of the whole sample, can be considered a medium-sized minority. Although this finding is not dissimilar to benchmarking data for the CORE-OM and Beck Depression Inventory when used within a large service (Barkham et al., 2001, p.195), it is an interesting phenomenon, worthy of further investigation.

Summary of section. In this section, I assessed what percentage of participants experienced reliable change, status change and clinically significant change, according to the overall change in their SI scores over the course of therapy.

- First, I calculated standardised metrics for this and future analyses, including aggregated data from previous studies, and identified a clinical significance cut-off score of 2.36 and RCI minimum values for two levels of significance: .97 ($p < .05$) and .64 ($p < .2$).

- Using the higher RCI standard (.97; $p < .05$) to analyse their pre-post change in scores, I found that 32% of participants experienced reliable change (31.1% improvement, 0.9% deterioration) and 68% of participants did not. Using the lower RCI standard (.64; $p < .2$), I found that 49.7% of participants experienced reliable change (46.6% improvement, 3.1% deterioration) and 50.3% of participants did not. I argued that the lower standard, equivalent to *clear and convincing evidence* (Stephen & Elliott, p.237), was more appropriate for use in clinical settings.
- Using the clinical significance cut-off score (2.36), I identified that the scores of 41.3% participants crossed over the threshold between clinical and non-clinical status, while the post-therapy scores of 58.7% of participants had the same status on the SI (i.e. clinical or non-clinical) as at pre-therapy.
- As a result, if applying the higher RCI standard, 26.2% of participants experienced clinically significant change over the course of therapy, while applying the lower RCI standard indicated that 33.3% of participants achieved clinically significant change by the end of therapy. These findings present a very different picture of change for individual participants than that suggested by the effect size (.95) for group pre-post scores identified in the previous section.
- Finally, by assessing change in this way, I observed that (a) there was no evidence of change over the course of therapy for approximately half of the participants (50.3%; RCI = .64; $p < .2$); and that (b) 20.4% of participants commenced therapy with scores in the non-clinical range on the measures and therefore unable to demonstrate clinically significant change, as defined by Jacobson and Truax (1991). This second result replicated the findings of Tashiro (2011). Both results raise questions about the shape of change that occurs in scores on the SI *during* the course of therapy, which will be examined in the next section of this chapter.

What is the Shape of Change in SI Scores over the course of Therapy (i.e. Linear, Nonlinear)?

In this section, I will report the results of a series of regression and other analyses, moving from assessing fit with a simple linear model to testing more complex nonlinear models, in order to investigate the *shape* of change in SI scores over the course of therapy.

First, through a simple regression analysis using pre-therapy and post-therapy scores only, I identified the linear trend within the data and distinguished those participants whose data may be unduly influencing this linear model, and others whose scores changed more or less than expected, according to the linear model.

Next, I will present the results of an investigation into the general pattern of change over the course of therapy when scores collected at mid-therapy data collection points were introduced. My purpose in this part of my study was to test my hypothesis that, at least for some clients, SI scores will decrease before they increase as participants develop their self-awareness in therapy and become more able to accurately discern their experience. First, I used a repeated measures ANOVA with a sub-sample of the data, then a multilevel linear analysis, working with the whole dataset, to assess if the data could fit a growth curve model.

Finally, I will present the results of a preliminary analysis of the shape of change at an individual level: first, I identified participants whose scores demonstrated reliable change (improvement or deterioration) at each data collection point; and, second, I investigated what happened to participants who continued in therapy beyond a data collection point at which their SI scores indicated reliable deterioration.

How well does pre-post change in SI scores fit a linear model? I conducted a simple regression analysis (Field, 2013, p.316-7) between participants' pre-therapy scores on the Strathclyde Inventory (the predictor variable) and their post-therapy scores (outcome) in order to identify a linear model that fits the data and

also to assess how well the model fits the data, known as the *goodness of fit* (Field, 2013, p.300). The process of regression transforms raw scores into residualised scores, removing the portion of the gain that could be predicted linearly from the pre-therapy score, so that participants with higher or lower than expected post-therapy scores can be identified. Figure 5.1 presents this model as a scatterplot that depicts each participant as a dot and inserts a line representing the linear model that fits best to the data:

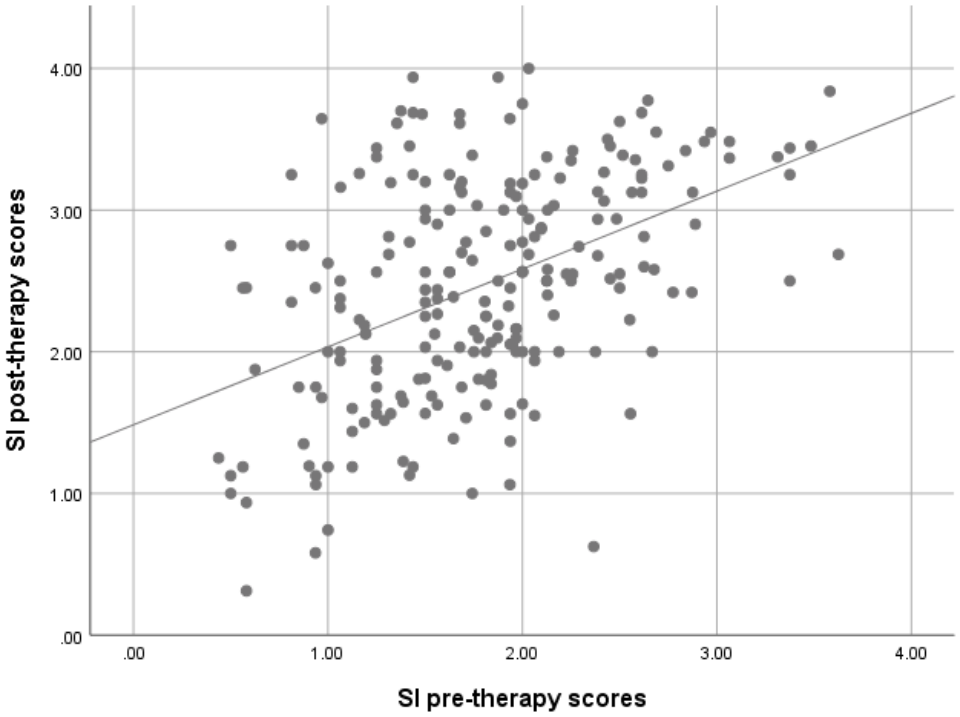


Figure 5.1. Scatterplot of the SI pre-therapy and post-therapy scores of 225 participants with fit line representing the general trend of change

An initial inspection of the scatterplot highlighted that, according to this dataset, the general trend for SI scores is to increase between pre-therapy and post-therapy. For example, according to the fit line (intercept = 1.49; slope = .55), a participant whose pre-therapy SI score was .50 would be expected to have a post-therapy score of 1.75. As Figure 5.1 demonstrates, while there are some participants whose pre-therapy and post-therapy scores fit the model precisely (i.e. their dot sits on the fit line), most of the 'real life' observed data vary from the model: for example, of the

three participants whose pre-therapy score was .50, two participants ended therapy with scores much lower than predicted by the model (1.00; 1.13) while the third had a post-therapy score that exceeded expectations (2.25).

The simple regression analysis assessed the goodness of fit of the model to the data. First, SPSS provided a summary of the model in which *r*, a simple correlation of the pre-therapy and post-therapy scores (predictor and outcome variables) = .46, which is a moderate correlation – much reduced from the test-retest correlation of .81, reported earlier in this chapter – indicating that there is a smaller association between pre-therapy scores and post-therapy scores than between pre-therapy scores and those taken before commencing the first therapy session. The analysis also identified r^2 = .212, which means that the pre-therapy scores account for 21.2% of the variation in post-therapy scores. The adjusted r^2 was .208, indicating minimal shrinkage (i.e. loss of predictive power) due to the large size of the sample.

Next, I conducted an ANOVA to produce an *F*-ratio. This is a ratio of the average variability that can be explained by the linear model and the average variability that is not explained by the model (Field, 2013, p.875). According to Field, a good model should have a large *F*-ratio, at least greater than 1, as a ratio of less than 1 would mean that there was more unsystematic than systematic variance in the data. In this case, the *F*-ratio was 59.9 ($p < .001$), which suggests that the regression model is a moderately good predictor of post-therapy scores.

The analysis also provided regression coefficients (Field, 2013, p.295) for the linear model. The first represents b_0, the intercept = 1.49 [95% confidence intervals: 1.22 (lower bound), 1.75 (higher bound)]. The second coefficient represents b_1, the slope = .55 [95% confidence intervals: .41 (lower bound), .69 (higher bound)]. Both *b*s are significantly different from 0 ($p < .001$). In summary, the results of these analyses mean that pre-therapy scores do make a contribution to post-therapy scores, to the degree represented by the gradient, but that the linear model

accounts for only 21.2% of the change that takes place in scores over the course of therapy.

Do the scores of some participants exert an undue influence over the linear model? I carried out some final checks to identify any cases that may be exerting undue influence over the model. I found no issues concerning *Cook's distance*, which is a statistic that measures the overall influence of a case on the model. There were no cases in my sample that had a Cook's distance value greater than 1. Next, I looked at *average leverage*, which analyses the influence of the observed value of the outcome variable over the predicted values. Field (2013, p.307) recommended using a cut-off score of no more than three times the average value, which for this sample, with one predictor, was = .027 (3 x.009). Three cases had values that fell outside this boundary: participants 163, 218, 570. Third, I examined values for *Mahalanobis distances*, which measure the distance of cases from the mean of the predictor variable. For this analysis, I used a cut-off value of 3.84, which is the chi-square critical value for one predictor at an alpha level of .05, and identified eleven problem cases: participants 391, 490, 561, 631, 589, 416, 475, 689, 163, 218, 570. Last, I checked the *covariance ratio (CVR)*, which measures whether a case influences the variance of the regression parameters. Ideally, the CVR should be as close as possible to 1. For a sample of this size with one predictor, an acceptable range can be established using the calculation of 1 plus/minus three times the average leverage value (.027, see above): i.e. .973 > CVR < 1.027. CVR values just outside this threshold are acceptable but others that are somewhat higher or lower may be problematic. In this case I identified one case with low CVR (participant 300) and six cases with high CVR (participants 589, 475, 689, 163, 218, 570).

Table 5.10 provides an overview of the potentially influential cases identified through this series of analyses:

Table 5.10. Overview of cases that may have undue influence on the model

Participants	Lev.	Mah.	CVR
163	x	x	x
218	x	x	x
300			x
391		x	
416		x	
475		x	x
490		x	
561		x	
570	x	x	x
589		x	x
631		x	
689		x	x

Notes. Lev. = leverage; Mah. = Mahalanobis distances; CVR = Covariance ratio.

These twelve cases represented 5.3% of the sample. Field (2013, p.347) recommended checking the Cook's distance for cases that cause concern. As in all cases the Cook's distance value is less than 1 then, according to Field, these individual cases are unlikely to have exerted undue influence in the model. To check this, I conducted a second simple linear regression removing these twelve cases from the sample. This analysis produced the following results (with original results in brackets): R = .435 (.46); R^2 = .190 (.212); adjusted R^2 = .186 (.208); F = 49.4 (59.9); b_0 (intercept) = 1.45 (1.49); b_1 (gradient) = .58 (.55).

These results indicated that the removal of the twelve potentially influential cases had little impact on the model and provided further support for an overall finding from these series of analyses that the linear model is a good fit for the data in the sample.

Which participants had scores that changed more or less than expected according to the linear model? Next, I used the casewise diagnostics function in SPSS to identify *outliers*, participants whose post-therapy scores had increased or decreased substantially more than predicted by the linear model: i.e. any cases that had a standardized residual that either exceeded 2 or was less than -2. According to Field (2013, p.345), in a normally distributed sample, 5% of cases would be

expected to fall outside these boundaries. This process highlighted eleven outliers (4.9%), almost exactly the result that would be expected by chance: six (participants 436, 76, 665, 387, 32 and 384) whose scores produced a standardised residual that was greater than 2, and therefore whose scores had increased substantially more than predicted when compared with other participants, and five (participants 264, 277, 316, 182 and 300) whose scores were lower than -2, and therefore whose scores had decreased significantly more than expected when compared with other participants. Only one of these eleven outliers (participant 300) had a standardised residual that was more than 3 or less than -3 and therefore could be described as an extreme case. Table 5.11 presents the results for these eleven participants:

Table 5.11. Outliers

Case Number	Std. Residual	SI post-therapy score	Predicted Value	Residual
436	2.40	3.94	2.28	1.66
76	2.35	3.65	2.02	1.63
665	2.11	3.70	2.24	1.46
387	2.06	3.94	2.52	1.42
32	2.02	4.00	2.60	1.40
384	2.04	3.69	2.28	1.41
264	-2.05	.58	2.00	-1.42
277	-2.08	1.00	2.44	-1.44
316	-2.15	1.06	2.55	-1.49
182	-2.16	.31	1.80	-1.49
300	-3.12	.63	2.79	-2.16

This result indicated that, at least according to their SI scores, the outcome of therapy for these eleven participants was significantly different from that of other participants and worth further investigation. Four of these participants: two 'improvers' (participants 665 and 32) and two 'deteriorators' (participants 300 and 316) were included in the case study series reported in Chapter 6.

What is the general pattern of change over time? In the preceding section of this chapter, I identified two interesting features of scores on the SI that should be explored further: (1) that a significant number of clients begin therapy with SI scores within the non-clinical range on the measure; (2) and that the scores of a

substantial proportion of participants did not demonstrate reliable change between pre-therapy and post-therapy. Although I have demonstrated in this section that a simple regression analysis can produce a linear model that fits well between pre-therapy and post-therapy scores, I was curious to know if the process of change during therapy, as measured on the SI, followed a linear course or if – as I suspected – some clients experienced a decrease in scores after they commenced therapy. If so, this would be consistent with my hypothesis that it is possible that some pre-therapy scores may be somewhat inaccurate, usually higher, due to a lack of self-awareness and that this initial inaccuracy may be 'corrected' as congruent functioning increases during the therapeutic process.

To test this hypothesis, I turned to the dataset that I had prepared containing scores for participants at the five main data collection points: from pre-therapy to 'after 40' sessions of counselling (N=714). I coded post-therapy scores at the data collection point closest to the number of sessions that the relevant participant had completed when they ended therapy (as reported in Table 5.2). First, I produced Figure 5.2, a boxplot of SI scores.

A visual inspection of Figure 5.2 confirmed that seven participants (participants 163, 218, 416, 475, 570, 589 and 689) were outliers at pre-therapy with scores that were significantly higher than the rest of the group. It was also clear that the bottom quartile of participants at the 'after 10' point stretches below the lower end of the 'pre-therapy' point, indicating that some participants' scores decreased after pre-therapy. A third point of interest in Figure 5.2, but less relevant for the current investigation, is that the scores of another three participants (182, 264 and 435) were outliers at the 'after 30' point, with scores falling well below those of the rest of the group. These scores represented post-therapy scores for two of the participants (182 and 264).

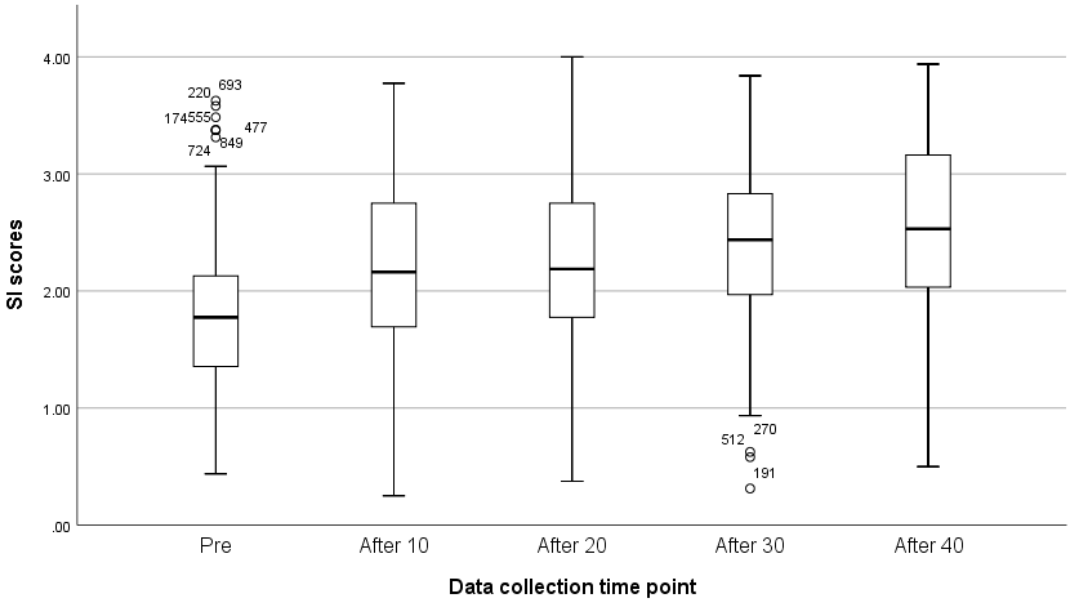


Figure 5.2. Boxplot of SI scores at five data collection points across therapy
Notes. Line within box = median. Box = middle 50% of observations. Whiskers = top and bottom quartiles. Participant codes for outliers at pre-therapy: 174 = 163; 220 = 218; 477 = 416; 555 = 475; 693 = 570; 724 = 589; 849 = 689. Participant codes for outliers at 'after 30' data collection point: 191 = 182 (post-therapy); 270 = 264 (post-therapy); 512 = 435.

Next I conducted a simple comparison of means at each data collection point: pre-therapy (N = 225; M = 1.79; SD = .65); after 10 sessions (N = 199; M = 2.23; SD = .71); after 20 sessions (N = 141; M = 2.25; SD = .72); after 30 sessions (N = 91; M = 2.35; SD = .75); and after 40 sessions (N = 58; M = 2.54; SD = .80). My primary goal in conducting this analysis was to obtain an eta-squared (η^2) effect size, as an estimate of non-linearity. The ANOVA demonstrated that the data collection point at which the data was collected had a significant effect on the SI scores, $F(4, 709) = 21.68$, $p < .001$, identifying a significant linear trend ($F(3, 709) = 71.15$, $p < .001$) and also a significant deviation from linearity ($F(3, 709) = 5.19$, $p < .01$). The measures of association were: $r = .299$; $r^2 = .089$; $\eta = .33$; $\eta^2 = .11$; providing statistical support for my hypothesis that there was likely to be a non-linear trend within the data collected during the course of therapy.

These results encouraged me to conduct a *repeated-measures ANOVA*. First I removed the ten outlying scores identified in Figure 5.2: seven at pre-therapy; three at the 'after 30' point. SPSS identified 39 participants in the remaining dataset who had data for all five data collection points so I used this sub-sample for the repeated-measures ANOVA. The mean scores at each data collection point for this sub-sample were: pre-therapy = 1.87 (SD = .55); after 10 sessions = 2.22 (SD = .61); after 20 sessions = 2.15 (SD = .59); after 30 sessions = 2.33 (SD = .66); and after 40 sessions = 2.62 (SD = .80). Figure 5.3 presents these mean SI scores across data collection points, with 95% confidence intervals and a line depicting the grand mean.

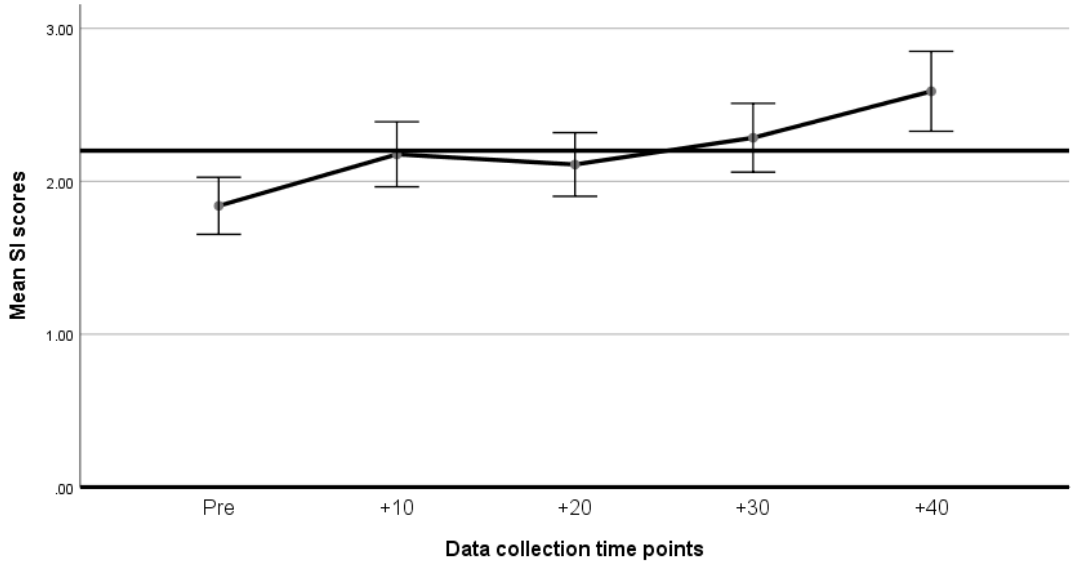


Figure 5.3. Line graph of mean SI scores across therapy for 39 participants who completed 40 sessions
Notes. Pre = pre-therapy; +10 = after 10 sessions; +20 = after 20 sessions; +30 = after 30 sessions; +40 = after 40 sessions. Error bars = 95% confidence intervals. Black horizontal line = grand mean.

Mauchly's test indicated that the assumption of sphericity (i.e. that there were equal variances across all levels of the repeated measures ANOVA) had been violated, $X^2(9) = 33.41$, $p = .000$. Therefore, I have reported Greenhouse-Geisser corrected tests ($\varepsilon = .68$). The results show that the data collection point had a

significant, moderate effect on participants' SI scores, $F(2.7,102.98) = 16.4$, $p < .001$, $\omega^2 = .07$.

Next, I tested within-subjects contrasts. This analysis identified that there was a significant difference in scores at each data collection point except between the 'after 10' point (level 2) and the 'after 20' point (level 3), $F = 1.18$, $p = .285$ (see Table 5.12).

Table 5.12. Tests of within-subjects contrasts between five data collection time points

Source	Data collection points	Type III Sum of Squares	*df*	Mean Square	*F*	Sig.	*d*
Data collection points	Pre vs. +10	4.786	1	4.786	18.736	.000	.60
	+10 vs. +20	.183	1	.183	1.175	.285	.12
	+20+ vs. +30	1.165	1	1.165	4.910	.033	.29
	+30 vs. +40	3.450	1	3.450	13.410	.001	.40
Error	Pre vs. +10	9.707	38	.255			
	+10 vs. +20	5.913	38	.156			
	+20 vs. +30	9.015	38	.237			
	+30 vs. +40	9.777	38	.257			

Notes. Pre = pre-therapy; +10 = after 10 sessions; +20 = after 20 sessions; +30 = after 30 sessions; +40 = after 40 sessions.

These results showed a nonlinear shape to the change in their SI scores across therapy for those 39 participants who accessed at least 40 sessions of therapy, comprising the following trend: first, after the first ten sessions of therapy, scores recorded a statistically significant change, indicating a medium to large effect; then, between ten and twenty sessions, scores recorded a non-significant change, with a small effect; next, between twenty and thirty sessions, scores recorded a statistically significant change, with a small to medium effect; and finally, during the last ten sessions, scores indicated a statistically significant change, with a medium effect. In other words, while these participants' aggregated scores clearly demonstrated improvement over time, a 'plateau' occurred between sessions 10 and 20. This plateau suggests that there may have been a variety of different trends emerging at that point in therapy in the data collected from this group of

participants: at its most simple, that some participants' scores increased, while others decreased, and that these opposing trends cancelled each other out when the data was aggregated. This theory of varied and contradictory patterns of change occurring mid-therapy may also be indicated, to a lesser degree, by the small to medium effect noted in the difference in scores between sessions 20 and 30. I will explore this idea further, still at a group level, in the next part of this section and then in more detail, at an individual level, in the final part of this section.

Does the change in SI scores over the course of therapy fit a nonlinear model? As described above, these results of the repeated measures ANOVA encouraged me to continue testing my hypothesis that change in SI scores over the course of therapy may fit a non-linear model. I used SPSS to conduct a *multilevel linear analysis* (Field, 2013, pp.814-866) to assess if the data could fit a growth curve model. Using a multilevel analysis enabled me to minimize the potential for statistical nonindependence by creating a hierarchy in the data in which each participant's scores, collected at two or more data collection points, were clustered by participant, enabling the analysis to recognize that there was a relationship between them. In addition, multilevel models are capable of handling datasets in which there is missing data (e.g. no SI score at a particular data collection point, either because the participant had completed therapy or due to an oversight); therefore it was possible for my whole dataset, encompassing scores across time for all 225 participants, to be used in this analysis. The analysis modelled a two-level hierarchical data structure in which participants were the level 2 variable and their score at each data collection point was the level 1 variable.

I carried out a series of analyses, testing the data first with a linear model (no points of change within the model), moving to a quadratic polynomial (one point of change within the model), then cubic polynomial (two points of change within the model), and finally a quartic polynomial (three points of change within the model). In order to keep the growth curve analysis simple, I allowed only the linear term to have a random intercept and slopes (Field, 2013, p.861-2). I was seeking to find the best fitting model by checking the change in the -2 log likelihood

(*-2LL*; the *deviance* statistic = -2 x log-likelihood; Field, 2013, p.763-4) score as the model became more complex. If the increase in the *-2LL* score in the more complex model, when compared with that of the previous model, was more than the relevant critical value for the chi-square statistic for the difference in the number of degrees of freedom in each model (i.e. if df_{change} = 1 then critical values for the chi-square statistic = 3.84 (p < .05) and 6.63 (p< .01)), then this means that the difference is significant and the model can be assumed to be a better fit. Table 5.13 contains the F-ratio, -2LL statistic and the degrees of freedom identified for each nonlinear model that I tested. The series of improvements in the *-2LL* statistic stopped at the quartic polynomial model when I found a non-significant difference.

Table 5.13. Growth curve

Model	*F-ratio*	*-2LL*	*df*	*-2LL* dif.
Linear	(1, 116.77) = 149.11, p < .001	1259.65	6	-
Quadratic	(1, 116.77) = 149.11, p < .001	1249.11	7	10.54**
Cubic	**(1, 354.62) = 12.10, p < .01**	**1237.19**	**8**	**11.92****
Quartic	(1, 351.00) = .794, p > .05	1236.40	9	0.79

Notes. *-2LL* = -2 x log likelihood score. *-2LL* dif = difference between-*2LL* score for this model and previous less complex model. ** = significant at p< .01 level. **Bold** = best fitting model.

The values for *Akaike's information criterion* (AIC), *Hurvich and Tsai's criterion* (AICC), *Bozdogan's criterion* (CAIC) and *Schwarz's Baysian criterion* (BIC) – alternative adjusted versions of the log-likelihood value (Field, 2013, p.825-6) - were also at their lowest level in the cubic polynomial model, confirming that this was the best-fitting approximation of the data.

The cubic polynomial model indicated that the data collection point significantly predicted SI scores, $F(1, 354.62) = 12.1$, p < .01. There was significant variance in intercepts across participants ($Var(u_{0j}) = 0.28$, SE = .04, p < .001), which confirmed that there was significant variation between participants' scores at pre-therapy. The analysis also identified significant variance between participants' slopes ($Var(u_{1j}) = 0.03$, SE = .01, p < .01), providing evidence that the change in individuals' scores over time varied significantly, pointing to heterogeneity in the

shape of client change on the SI over time. However, the analysis found no significant covariance of slopes and intercepts when evaluated with an autoregressive covariance structure, which assumes that variances will be heterogeneous (ARH1 rho = -.06, SD = .14, $p < .7$), nor with an unstructured alternative (Cov(u_{0j}, u_{1j}) = -.01, SD = .01, $p < .7$).

By investigating the dataset using multilevel linear modelling, I have demonstrated that the shape of change in SI data for this participant group fits a cubic polynomial model, which means that two points of change across therapy can be detected. This is consistent with the finding of the repeated-measures ANOVA using data from the sub-sample of 39 participants who had a full set of data across forty sessions of therapy, for whom there was a general positive trend in the early stages of therapy, followed by a plateau in mid-therapy, and finally a second upward turn in scores in the final phase of therapy. However the multilevel linear analysis also confirms the variation in experience for participants within the group: that there is significant variety in the level of congruent functioning reported by participants at pre-therapy, significant differences in the way in which their functioning develops over time in therapy, and no significant relationship between participants' pre-therapy SI scores and the subsequent process of change in congruent functioning that they experienced.

What proportion of participants have scores that reliably increased or decreased between data collection points during therapy? As a final step, I wanted to investigate change in scores for participants at an individual level that occurred between data collection points across therapy. I was interested in finding out the pattern of reliable change that took place between data collection points. More specifically, I anticipated that this would identify if there were participants whose scores had reliably deteriorated after they started therapy but who had not gone on to record reliable deterioration in their overall pre-post change score. First, I calculated the proportion of participants whose scores increased or decreased by at least the amount of change identified by the RCI (.64, $p < .2$). I chose this lower standard as it seemed an appropriate level of probability for the reason outlined

earlier in this chapter. I used SPSS to calculate the difference between each individual participant's scores on each consecutive pair of data collection points: pre-therapy, after 10 sessions, after 20 sessions, after 30 sessions, and after 40 sessions. If there was no data available for any time point, either due to data collection error or because the participant had completed therapy, this calculation returned a missing data count. The results of this analysis are presented in Table 5.14.

Table 5.14. Reliable improvement or deterioration between data collection time points

		RCI ↑	RCI ↓	No change	Total N
Pre - +10	N([a])	60 (33)	4 (3)	135	199
	%	30.2	2.0	67.8	
+10 - +20	N([a])	24 (18)	6 (3)	90	120
	%	20.0	5.0	75.0	
+20 - +30	N([a])	13 (5)	3 (1)	60	76
	%	17.1	3.9	78.9	
+30 - +40	N([a])	11 (11)	2 (1)	40	53
	%	20.8	3.8	75.4	

Notes. Pre = pre-therapy; +10 = after 10 sessions; +20 = after 20 sessions; +30 = after 30 sessions; +40 = after 40 sessions. RCI = .56, $p < .2$; RCI ↑ = reliable improvement; RCI ↓ = reliable deterioration. [a] = number of participants in group who ended therapy at this point.

Table 5.14 demonstrates that the largest proportion of participants experienced a reliable improvement in their SI scores between pre-therapy and after 10 sessions. For 33 of these 60 participants (55%) this was also the point at which they ended therapy. In contrast, this was the point at which the smallest proportion of participants whose scores reliably deteriorated (N=4; 2%) was recorded. For 75% of this group (N=3), this was also the moment at which they ended therapy.

The proportion of participants whose SI scores suggested reliable improvement decreased across the next two data collection points (20% after 20 sessions then 17.1% after 30 sessions) then increased slightly to 20.8% after 40 sessions, the point at which most participants in the sample ended therapy. I returned to the data and checked how many of the participants whose data

demonstrated a reliable improvement ended their therapy at that point and found that this was true for 18 of 24 participants (75%) whose scores reliably improved between 10 and 20 sessions, 5 of 13 participants (38.5%) whose scores reliably improved between 20 and 30 sessions, and all of the 11 participants whose scores reliably improved between 30 and 40 sessions. This pattern was mirrored in the data for participants whose scores demonstrated reliable deterioration: I found that 3 of 6 participants (50%) whose scores reliably deteriorated between 10 and 20 sessions ended therapy at that point, as did 1 of 3 participants (33.3%) whose scores reliably deteriorated between 20 and 30 sessions, and 1 of 2 participants (50%) whose scores reliably deteriorated between 30 and 40 sessions. Indeed, it was only after session 30, that the majority of participants whose scores reliably changed (improvers = 61.5%; deteriorators = 66.6%) continued in therapy, suggesting that at this point, whether improving or deteriorating, these participants were motivated to complete the full number of sessions available to them. The preliminary evidence offered by this rough analysis suggests that there may be a relationship between reliable improvement in SI scores and the decision to end therapy, which should be investigated in future studies.

What happened to participants who continued in therapy after a data collection point at which SI score suggested reliable deterioration? I was curious to discover what happened to participants whose scores indicated reliable deterioration during therapy and continued with the process. As Table 5.14 indicated, there were seven participants who fell into this category: participant 669, whose score reliably deteriorated after 10 sessions in therapy; participants 314, 471 and 484, whose scores reliably deteriorated after 20 sessions; participants 139 and 300, whose scores reliably deteriorated after 30 sessions and, for participant 300, again at the end of therapy, after 50 sessions; and participant 441, whose score reliably deteriorated after 40 sessions. Table 5.15 presents an overview of the change in their scores at each data collection point across therapy, along with their pre-therapy and post-therapy scores and overall pre-post change in score. Difference in scores identified as reliable change (.64; $p < .2$), whether improvement

or deterioration, is highlighted in bold; pre-post change that can be defined as clinically significant change (Jacobson & Truax, 1991), in which the participant's scores moved from the clinical range on the measure to the non-clinical range by the end of therapy, is identified with an asterisk.

Table 5.15. Overview of change in SI scores across therapy for participants who continued in therapy after reliable deterioration

Client	Pre	+10dif	+20dif	+30dif	+40dif	+50dif	Post	P-Pdif
669	2.00	**-0.65**	0.28	-	-	-	1.63	-0.37
314	2.61	**0.94**	**-0.74**	0.50	0.32	0.06	3.69	**1.08**
471	1.25	**1.56**	**-0.75**	0.38	**0.94**	-	3.38	**2.13***
484	2.38	0.00	**-1.13**	0.06	**0.69**	-	2.00	-0.38
139	1.42	0.35	0.52	**-0.77**	-0.39	-	1.13	-0.29
300	2.37	-0.43	0.09	**-1.09**	0.58	**-0.89**	0.63	**-1.74**
441	0.81	[miss]	0.07	**1.62**	**-2.00**	**2.75**	3.25	**2.44***

Notes. Pre = pre-therapy SI scores; +10dif = difference in SI scores between pre-therapy and 10 sessions data collection point; +20dif = difference in SI scores between 10 and 20 sessions data collection points; +30dif = difference in SI scores between 20 and 30 sessions data collection points; +40dif = difference in SI scores between 30 and 40 sessions data collection points; +50dif = difference in SI scores between 40 and 50 sessions data collection points; Post = post-therapy score; P-Pdif = change in scores between pre- and post-therapy. [miss] = missing data; - = no data (client had ended therapy). **Bold** = reliable change. * = clinically significant change.

As Table 5.15 shows, the pattern of change in SI scores across therapy for these seven participants varied considerably. Participant 669 continued in counselling for a further ten sessions. Their final SI score was not reliably different from their pre-therapy score (pre = 2.00; post = 1.63). For participants 314 and 471, the reliable deterioration in their scores after 20 sessions had followed reliable improvement in their scores at the 'after 10 sessions' data collection point. After this drop, participant 314's scores continued to increase at a non-significant rate for the remainder of their therapy (beyond the usual 40 sessions limit), resulting in reliable improvement in scores overall (1.08). This was not clinically significant change as this client's pre-therapy score was in the non-clinical range on the measure (2.61). In contrast, participant 471, whose pre-therapy score was in the clinical range (1.25), experienced a further reliable improvement of scores between sessions 30

and 40 (.94), resulting in scores indicating clinically significant change at the end of therapy (post-therapy score = 3.38; pre-post change = 2.13).

Participants 484 and 139 both ended therapy with SI scores that had not reliably changed from those recorded at pre-therapy (484 pre-post change = -0.38; 139 pre-post score = -0.29) but via different pathways. Following the reliable deterioration in their score after session 20 (-1.13), participant 484's scores remained at a similar level between sessions 20 and 30, then reliably improved between sessions 30 and 40 (0.69), when they ended therapy. In contrast, participant 139's scores, which had gradually but non-significantly increased between pre-therapy and session 20, reliably deteriorated between sessions 20 and 30 (-0.77), then continued to decrease, although to a non-significant degree (-0.39), by the end of therapy.

The final two participants in this group had vastly different experiences in therapy, according to the change in their SI scores across the process. Participant 300 began with a score just into the non-clinical range (2.37), whereas participant 441's pre-therapy score was at the lower end of the clinical range (.81). A critical point in therapy for both participants, according to their SI scores, was the period between sessions 20 and 30. At the 'after 30 sessions' data collection point, participant 300's score indicated reliable deterioration (-1.09), while participant 441's score reliably improved (1.62). From this point onwards, the scores of both participants fluctuated between data collection points: participant 300's scores indicated non-reliable improvement (0.58) by session 40, followed by a second reliable deterioration by the end of therapy (-.89); on the other hand, participant 441's scores showed reliable deterioration by session 40 (-2.00) then ended therapy after 50 sessions with another change that indicated reliable improvement (2.75).

Participant 300 was the only one of this group of seven whose scores reliably deteriorated across the course of therapy. Participants 314, 441 and 471, despite their SI scores reliably deteriorating during the course of therapy, ended with reliably improved scores, which for participants 441 and 471, could be defined as clinically significant change.

Figure 5.4 displays the individual pathways of each participant over the course of therapy, according to their SI scores.

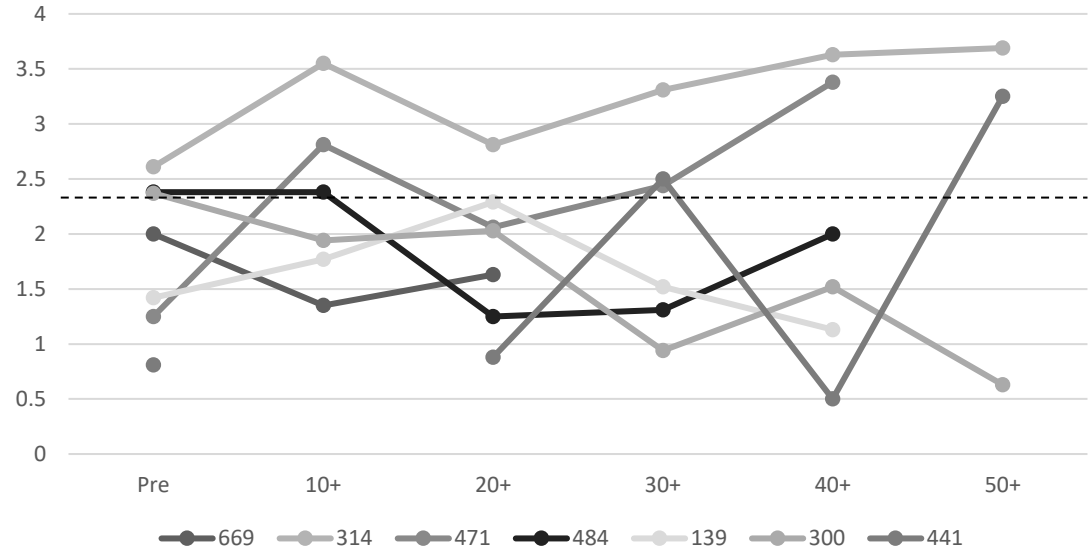


Figure 5.4. Line chart depicting SI scores at data collection time points for participants who continued in therapy after reliable deterioration
Notes. Dash line indicates clinical cut-off score (2.36). Participant 441 has no line between Pre and 20+ because of missing data at 10+ data collection time point.

At the beginning of this chapter section, I had indicated my interest in the experience of (1) participants who recorded pre-therapy SI scores that were within the non-clinical range and also (2) participants whose outcome indicated no reliable pre-post change. This small sub-sample of participants provided some preliminary answers to my questions.

First, three participants (300, 314 and 484) commenced counselling with scores in the non-clinical range, although the scores of two of these participants (300 and 484) were only just over the threshold. My hypothesis was that, for these clients, their pre-therapy scores might be inaccurately high due to low congruent functioning (i.e. low self-awareness) and that their process in therapy might include a decrease in scores at some point, as their congruent functioning increased and they became more self-aware. Figure 5.4 demonstrates that, having both recorded reliable deterioration in their SI scores after session 20, the scores for participants

314 and 484 began to rise and the cumulative increase in their scores after session 20 until the end of therapy equated to reliable improvement: for participant 314, a total increase of 0.88 (0.50 + 0.32 + 0.06); for participant 484, a total increase of 0.75 (.06 + 0.69). Therefore, although the post-therapy scores of these two participants did not return to their pre-therapy levels, their scores did indicate reliable improvement following reliable deterioration, which may provide support for my hypothesis. In contrast, participant 300's scores continued to deteriorate. The experience in therapy of this participant was examined as a case study contributing to the final part of my investigation, which will be presented in Chapter 6.

The experience of three participants (139, 484 and 669), whose pre-post change was too small to be considered reliable, is more varied. While the scores of participant 484 recorded reliable deterioration then reliable improvement by the end of therapy, participant 139's scores followed the opposite pattern: accumulative reliable improvement between pre-therapy and session 20 (0.87; 0.35 + .52) followed by reliable deterioration (-.77) between sessions 20 and 30 and indication of a further, non-significant, decrease (-.39) at the end of therapy, suggesting a slowing rate of deterioration. Given the varied pattern of scores depicted in this small sample of participants, it is impossible to predict what might have occurred for participant 669, whose scores reliably deteriorated during the first ten sessions of therapy then recorded a small, non-significant increase by session 20, if they had continued in therapy beyond this point.

However, what is clear when looking at SI scores *during* therapy, not just between the beginning and end of the process, is that change measurable on the SI is taking place. It may take place slowly, perhaps for some incrementally, unable to be captured between ten sessions but evident when comparing scores across twenty sessions, as in the case of participants 314, 484 and 139. It may be offset by an earlier decrease in scores during which, I propose, the therapeutic process enables clients to become more accurately aware of their own experience, and for some can result in an apparent lack of change in scores between the beginning and

end of therapy. A final conclusion from this examination of the pattern of change in SI scores for participants in this small sub-sample is that this decrease in scores can occur at any time in the process: early in therapy (e.g. participant 669), in the middle (e.g. participants 314, 471, 484, 139 and 300) and towards the end, or at the end of therapy (e.g. participants 441 and 300) and, as the data for these participants demonstrates, does not necessarily determine a negative outcome of therapy.

Summary of section. In this section I investigated the shape of change in scores on the Strathclyde Inventory over the course of therapy, first through the simple lens of a linear model, next by testing the data with non-linear models, and finally by examining the data of individual participants, identified by calculating if their scores changed to a reliable degree between data collection time points. In doing so, I was able to investigate more closely the experience of some participants who scored highly on the SI at pre-therapy and others whose pre-post change in scores suggested that no significant change had occurred. The main findings in this section are:

- A linear model is a good fit for the data: $F(1,223) = 59.9$, $p < .001$, $r^2 = .212$. Although twelve potentially influential cases within the sample (N=225) were identified, tests demonstrated that these had little impact on the model. Eleven 'outliers' were identified whose standardized residual scores were greater than 2, or less than -2 and therefore whose pre-post change on the SI was much greater or less than predicted.
- However, a cubic polynomial model is a better fit for the data: $F(1, 354.62) = 12.1$, $p < .01$, $-2LL = 1237.19$. Estimates of the covariance parameters indicated that there was significant variation amongst participants in, and no relationship between, pre-therapy SI scores and the shape in which their congruent functioning developed over the course of therapy. These findings demonstrate that, although there is a general trend of gradual improvement in scores between the beginning and end of therapy, there is a high degree

of variation in the way in which change takes place for individual participants.

- A minority of participants recorded reliable change (.64; $p < .2$) between data collection points: 21.2 - 32.2%. I found preliminary evidence of a relationship between reliable change in SI scores between data collection points and the decision to end therapy.
- Of the seven participants who continued in therapy following a data collection point at which a reliable deterioration in their SI score was recorded, only one went on to experience overall reliable deterioration in their SI scores. In contrast, three participants recorded reliable improvement overall.
- Looking at the change of scores for this small sub-sample enabled some investigation of the shape of change experienced by participants with (1) non-clinical pre-therapy scores, and (2) no reliable pre-post change. Three of these seven participants had pre-therapy scores in the non-clinical range. The subsequent change in scores for two of these participants demonstrated cumulative reliable improvement, although the overall pre-post change for one of these two participants was not significant. The scores of the third participant in this group recorded overall reliable deterioration at the end of therapy.
- A different configuration of three participants in this sub-sample did not achieve pre-post reliable change, with cumulative improvement in scores offsetting the deterioration that had taken place. For two participants, improvement followed deterioration; the scores of the third participant followed the opposite pattern.

In What Ways do Clients' Scores on Individual SI items Change over the course of Therapy?

The final way that I chose to investigate change in SI scores over the course of therapy was to examine the change that takes place on scores for individual items. First, I examined the sensitivity to change indicated for each item, then as a

result of my findings, I returned to Rasch measurement to test if there was any evidence that the function of particular items changed between pre-therapy and post-therapy, according to participants' scores.

Are some items within the SI more sensitive to change? I used t-tests to check the statistical significance of any difference between aggregated mean pre-therapy and post-therapy scores for each item in the 20-item version of the SI developed in my first study. As the 20-item version includes four items that were featured in the SI-31 but not the SI-16 there is a lower N for those items as only data from those SI-31s in the sample could be used in the analyses. Table 5.16 presents the results.

Table 5.16. Item sensitivity to change

Item	Pre-therapy			Post-therapy			*t*	*d*
	N	M	SD	N	M	SD		
1	225	1.59	1.07	225	2.16	1.10	-5.63**	.53
2	224	1.53	1.08	224	2.38	1.11	-8.17**	.77
3	225	1.76	1.18	224	2.55	1.16	-7.21**	.68
4	224	1.98	1.02	225	2.59	1.04	-6.23**	.59
5	224	1.77	1.19	224	2.27	1.24	-4.27**	.40
6	225	1.83	1.22	224	2.63	1.20	-6.94**	.65
7	225	1.40	1.10	224	2.24	1.13	-7.97**	.75
8	224	2.11	1.12	224	2.66	1.12	-5.19**	.49
9	225	1.24	1.03	225	2.12	1.17	-8.38**	.79
10	225	2.59	1.02	225	2.84	.97	-2.70*	.25
11	224	1.79	1.08	225	2.47	1.00	-6.96**	.66
12[a]	104	2.43	1.27	123	3.04	1.07	-3.91**	.52
13	223	1.11	1.01	224	1.91	1.12	-7.92**	.75
14[a]	106	2.89	.87	121	3.07	.79	-1.62	.22
15	225	1.40	1.13	223	2.16	1.16	-7.07**	.44
16	224	1.85	1.10	223	1.57	1.10	-6.93**	-.25
17	225	1.43	.98	222	2.29	1.05	-8.98**	.85
18[a]	105	2.46	1.07	120	3.18	.93	-5.39**	.72
19[a]	104	1.91	1.08	121	2.64	1.12	-4.97**	.66
20	223	1.74	1.17	223	2.58	1.17	-7.56**	.72

Notes. See Appendix D for full list of items in SI-20. [a] = item not included in SI-16. ** = $p < .001$; * = $p < .01$. Cohen's *d*: .20 = small effect; .50 = medium effect; .80 = large effect.

As Table 5.16 shows, there was a statistically significant difference at the $p < .001$ level between pre-therapy and post-therapy means on 18 of the 20 items. For item 10 (*I have made choices based on my own internal sense of what is right*), the difference was statistically significant but at a lower standard ($p < .01$). There was no significant difference between the pre-therapy and post-therapy means for item 14 (*I have been aware of my feelings*).

There are six items on which pre-post scores demonstrated a large effect ($d > .72$; max. = .85): item 17 (*I have been able to resolve conflicts within myself*); item 9 (*I have found myself 'on guard' when relating to others*); item 2 (*I have condemned myself for my attitudes or behaviour*); item 7 (*I have looked to others for approval or disapproval*); item 18 (*I have felt myself doing things that are out of character for me*); and item 20 (*I have felt it is all right to be the kind of person that I am*). This result suggests that these items are the most sensitive to the type of change that participants experienced over the course of their therapy. Remembering that items 9, 2, 7 and 18 were negatively worded and therefore reverse scored, this means that, over the course of therapy, participants noticed a large positive change in their ability to know, understand, trust and accept themselves in the ways that were highlighted by these particular items.

There were three items with small effects ($d < .25$). These included items 10 and 14, already identified as having less or no significant difference between means at pre-therapy and post-therapy, and also item 16 (*I have felt true to myself*). The effect size for item 16 was both small and negative (-.25), a striking result. I double-checked the data to ensure that there was no error in my dataset that would cause this result. I found none. Therefore, this result raised an interesting question: why might participants feel that they were less true to themselves by the end of therapy? One possibility was that the meaning of the item had changed for them over the course of therapy. As a result, I decided to investigate whether it was possible that the meaning of this and other items may have changed for participants between the start and end of therapy.

Do any SI items change in meaning for clients between pre- and post-therapy? Rasch measurement has two functions that can be used to compare the performance of instruments when completed by two different groups: Differential Test Functioning (DTF) and Differential Item Functioning (DIF). DTF compares the performance of the test as whole, whereas DIF allows for the comparison of individual *item difficulty* between groups (Linacre, 2012) indicated by different positioning within the hierarchy of item difficulty (as discussed in Chapter 4). Typically, DIF has been used to assess gender bias in instruments or to compare difference in meaning of items across national groups (Bond & Fox, 2015, p.106).

In this study, I decided to use first DTF and then DIF to assess if there are any differences in the way that the SI and its individual items functioned at pre-therapy and post-therapy for the participants in my dataset. This was a simpler method than the multilevel longitudinal Rasch measurement model used by Pastor and Beretvas (2006); more appropriate for my level of knowledge and proficiency with Rasch. First, I used the Rasch program, Winsteps, to conduct a DTF. Linacre (2012) argued that DTF is a useful approach because it provides a better overall sense of test bias than focusing on individual items. In DTF, two separate analyses are carried out, first on the pre-therapy data, second on the post-therapy data, producing item difficulty measures for each group and these results are brought together in a cross-plot of items. Figure 5.5 presents the results of the DTF.

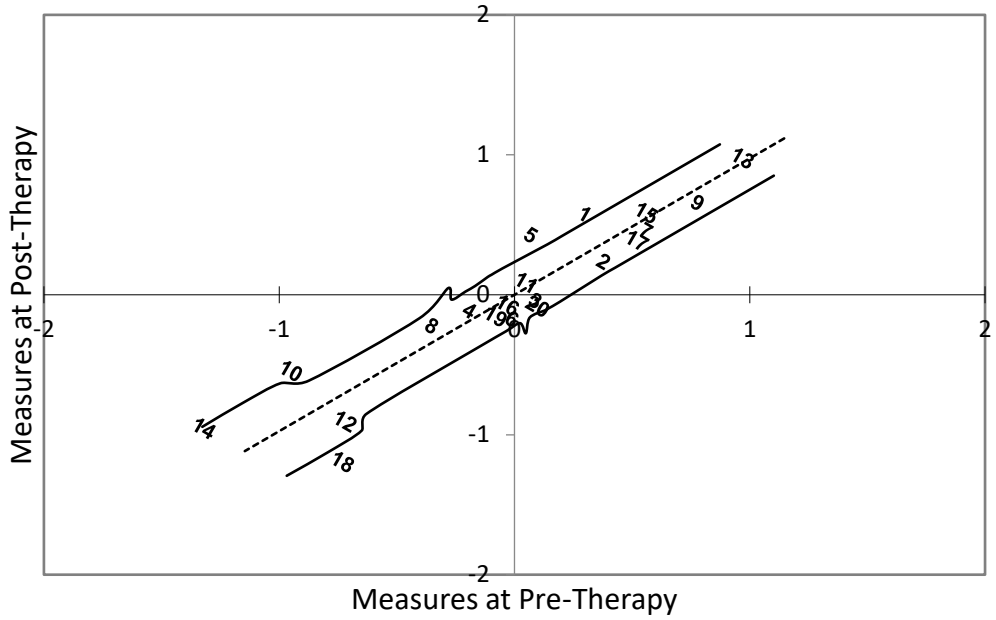


Figure 5.5. Differential test functioning plot comparing pre- and post-therapy
Lines = approximate 95% confidence bands. Dashed line = trend line through mean of both sets of items. Numbers within plot = item numbers.

The x-axis of the plot displays the item difficulties for the pre-therapy group and the y-axis displays the item difficulties for the post-therapy group. The dashed line is a trend line through the mean of both sets of items. The black lines represent approximate 95% confidence bands. Items are represented within the plot by the number of the item. Items outside the lower confidence band were more difficult for participants at pre-therapy than at post-therapy. There is one item identified by the DTF in this category: item 18 (*I have felt myself doing things that are out of character for me*). This indicated that participants found it 'easier' to respond to this item at the end of therapy.

Items plotted outside or just on the upper confidence band were more difficult for participants at post-therapy. There were four items in this category: item 1 (*I have been able to be spontaneous*), item 5 (*I have experienced very satisfying personal relationships*), item 10 (*I have made choices based on my own internal sense of what is right*) and item 14 (*I have been aware of my feelings*). This

result suggested that participants found it harder to respond to these items at the end of therapy.

Next, I used Winsteps to carry out a DIF. Linacre (2012) noted that this is the more common approach used to investigate item functioning because the analysis is conducted as one combined analysis in which only one thing is altered, in this case the point in time at which the person completes the measure, so that any effect this produces can be detected. The results of the DIF were produced in the form of tables and graphs. including a Person DIF plot that compared the items according to their difficulty for the group of participants at each data collection point, measured in logits. This plot is presented as Figure 5.6. The difficulty measures for each item at pre-therapy are displayed as black diamonds connected by a black line. The difficulty measures on the same items at post-therapy are displayed as grey squares connected by a grey line.

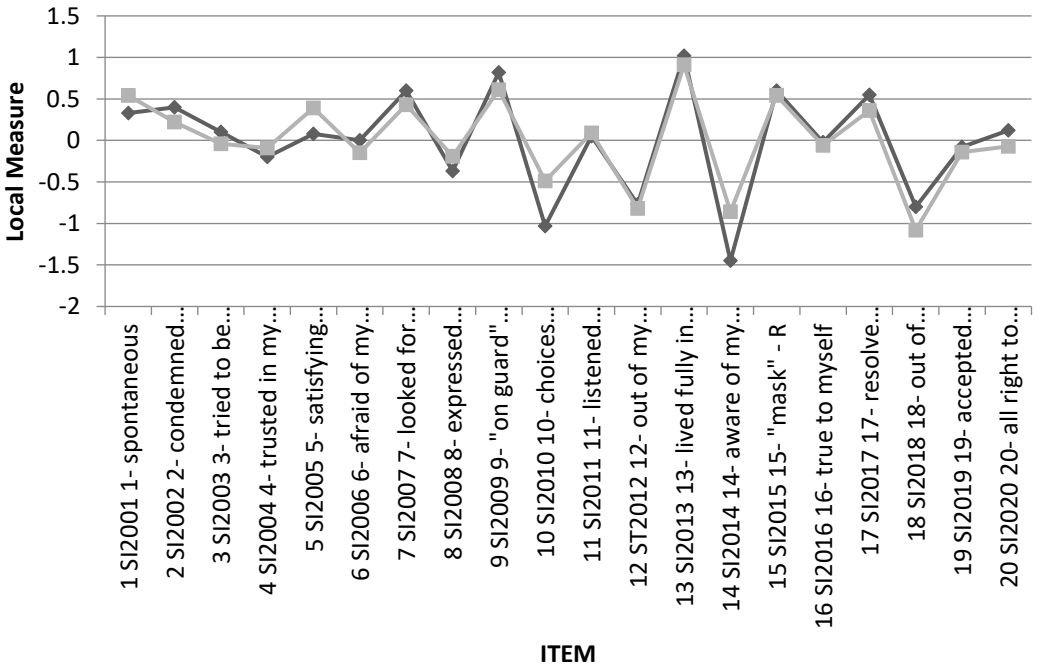


Figure 5.6. Person DIF plot contrasting t-values per item at pre- and post-therapy
Notes. Line 1(black) = pre-therapy; Line 2 (grey) = post-therapy.

A visual inspection of Figure 5.6 identified that the items appearing to have the largest differences between pre-therapy and post-therapy were items 5, 10 and 14, three of the items that had been highlighted by the DTF. I checked the statistical significance of these differences. Winsteps uses Welch's t-statistic to test statistical significance in DIF. According to the information produced (Winsteps output table 30.1), the same three items were the only items with a statistically significant difference between item difficulty at pre-therapy and post-therapy: item 5 (t = 2.84, p < .01); item 10 (t = 4.53, p < .0001); item 14 (t = 3.34, p < .01). However, only two of these items had a DIF contrast greater than .5 logits, the minimum size recommended by Linacre (2011) to merit further investigation: item 5 = .32; item 10 = .53; item 14 = .59. Linacre (2012) noted that the DIF impact depends on the length of the test and that therefore a small DIF size can be found to be highly significant while a large DIF size may be reported as not statistically significant. In both cases, the items were more difficult for participants at post-therapy.

In summary, the DTF highlighted potential change in the functioning between pre-therapy and post-therapy of five items. By carrying out the DIF, this result pinpointed two of these items: item 10 (*I have made choices based on my own internal sense of what is right*) and item 14 (*I have been aware of my feelings*). The analyses indicated that participants found responses to both items more difficult to endorse, relative to other items, at the end of therapy than at the beginning. As Pastor and Beretvas (2006, p.116) noted, trends of this nature infer a change in the way that the item is functioning and require consideration of what this may mean. In this case, this result suggests a change in participants' perception of each item. This could be that their understanding of the meaning of the item changed over the course of therapy (e.g. what it *really* means to make choices based on an internal sense of what is right), or alternatively that their understanding of its relevance to their own experience changed (e.g. have developed greater appreciation of how challenging it is for them to recognise and act upon their internal sense of what is right). Both of these possible explanations

suggest some degree of increased self-awareness, an indicator of increased congruent functioning.

Interestingly the results of both the DTF and DIF found no change in functioning of item 16 (*I have felt true to myself*), the item that originally prompted this investigation into the comparative performance of individual items at pre- and post-therapy. This suggests that a literal interpretation of this result could be made: that participants did indeed feel less true to themselves by the end of therapy. The mean post-score on this item is 1.57, significantly lower that the clinical significance cut-off score (2.36) and the lowest individual item mean score at post-therapy; lower even than item 13 – *I have lived fully in each new moment* – which my analyses in Chapter 4 identified as having the highest item difficulty in the instrument. Paradoxically, I see this result as an indicator that participants have indeed increased their congruent functioning by the end of therapy: that they are now more aware of their incongruence than they were at the beginning and can now recognise the existence of a gap between their current way of living in the world and their deeply-felt needs.

Summary of section. In this section, I investigated change over the course of therapy as indicated by the individual items that comprise the SI (20 item version). The main findings were:

- There was strong evidence of pre-post change when examining difference in mean scores on individual SI items. Eighteen of the twenty items demonstrated statistically significant change at $p < .001$ level. Effect sizes ranged from large (6 items; $d > .72$; max. = .85) to small (3 items; $d < .25$).
- One item (16 - *I have felt true to myself*) had a negative small effect size ($d < -.25$), indicating that participants' scores on this item tended to deteriorate by the end of therapy. This result was investigated using Rasch measurement (differential test functioning and differential item functioning). No change in item functioning was detected, confirming that a literal interpretation of this result may be assumed. I have proposed that

this pre-post change in scores on this item is an indicator of participants' increased congruent functioning by the end of therapy.

- I identified two other items (10 - *I have made choices based on my own internal sense of what is right*; and 14 - *I have been aware of my feelings*) as ones that could be perceived as potentially problematic. These were the two items for which pre-post change in scores had either less statistical significance (item 10; $p < .01$) or not statistically significant (item 14) and demonstrated least sensitivity to change (item 10, $d < .25$; item 14, $d < .22$). Further analyses - first, differential test functioning, then differential item functioning – identified these two items as becoming more difficult for participants to score by the end of therapy, possibly as a result of increased congruent functioning.

Chapter Summary

The results of this study have provided evidence that the SI is sensitive to change in scores over the course of therapy, an important requirement for an instrument designed to measure the outcome of therapy. In doing so, I have identified some interesting complexities in the findings that raise further questions.

First, the size of pre-post change in the aggregated SI scores for participants in this study was .95, a large effect size that indicates that mean scores on the measure changed by almost one standard deviation between the beginning and end of therapy. However, I found a more modest result when analyzing the pre-post scores at an individual level using Jacobson and Truax's (1991) approach to assessing clinically significant change. From this perspective, I identified that, using my recommended $p < .2$ standard, around one third (33.3%) of participants 'recovered' (their pre-post change in scores indicated clinically significant change), while under half (46.6%) of participants 'improved' (their pre-post change in scores indicated reliable change in the direction of improvement on the measure). A very small minority of participants (3.1%) ended therapy with pre-post change in SI scores that indicated a reliable deterioration. According to Lambert (2013, p.178), it is typical to find that the effect size statistic overestimates the proportion of

individuals whose scores can be defined as clinically meaningful change. The discrepancy between measurement of change according to effect size in comparison to clinically significant change, confirms that using effect size as a metric for assessing efficacy and effectiveness within counselling and psychotherapy research misrepresents the experience of individual clients.

A more conservative estimate of change, the clinical significance results in this study are consistent with the findings of other studies using Jacobson and Truax's approach: for example, Eisen, Ranganathan, Seal & Spiro (2007), who compared change measured by effect size, standard error of measurement and reliable change index (RCI), and found that the RCI method identified the smallest number of improved individuals and the highest number of individuals showing no change. Indeed, Bauer, Lambert & Nielsen (2004) recommended the Jacobson and Truax method as providing the most moderate estimate of reliable change when compared with four alternative methods for calculating reliable change. Their results indicated that the typical distribution of results when testing the same sample analysed with these different methods was: recovered (11.9-21.2%), improved but not recovered (12.7-24.9%), unchanged (42.0 – 66.6%), and deteriorated (3.1-8.8%). The results of this study, whether assessed using the lower or higher RCI standard, indicate a reasonable, indeed somewhat favourable, match with these distributions: a greater proportion of participants experienced clinically significant change, a similar proportion of participants (if combining recovered and improved but not recovered) recorded reliable change, and a low, or lower, proportion of participants had scores that indicated deterioration.

These results, being well within acceptable expectations for change, measured from these two different perspectives, confirm that the SI can demonstrate an appropriate sensitivity to change in scores over the course of therapy that is equivalent to other outcome measures. Indeed, given that other instruments tested in this way tend to be measures of distress, these results indicate that the SI may be capturing a higher degree of change than is typical for measures of experiential functioning (Elliott, 2001).

However, questions remain: what happened over the course of therapy for the 50.3% of participants in this study whose scores did not change by the minimum value required to indicate reliable change, and the 20.4% of participants who began therapy with SI scores in the non-clinical range on the measure? As this study has demonstrated via the patterns of change indicated for the sub-sample of seven participants selected for further examination, finding no change between pre- and post-scores does not mean no change occurred during therapy. The experience in therapy for the majority of participants, whose pre-post scores on the SI suggested no change or cannot be assessed for clinical significance because the pre-therapy score was in the non-clinical range, requires further investigation.

Next, this study found change in scores could fit a linear model describing a gradual improvement across time but no direct evidence to support the *dose effect* model of change (Howard et al., 1986) in which client improvement is associated with length of time in counselling: there was no statistically significant correlation between pre-post difference in SI scores and number of sessions accessed. However, there was some preliminary evidence of a possible relationship between participants experiencing a reliable change in scores at a particular data collection point and the ending of therapy at that time that may support the *good enough level of improvement* model (Barkham et al., 2006) in which participants leave therapy when they have improved to a degree that is perceived to be 'good enough'. However, it is still unclear whether this experience of change leads to the decision to end or if the change in score is influenced by a decision to end - a *flight into health*, which has been traditionally viewed, especially from a psychoanalytic perspective, as an 'escape tactic' (Frick, 1999). Further research is required to investigate this relationship in more detail.

The phenomenon of early change (e.g. Lutz et al., 2009; Stulz et al., 2007) was also visible within this study: 30.2% of participants recorded SI scores that indicated reliable improvement after 10 sessions: 55% of these participants ended therapy at that data collection point; this study did not investigate the eventual outcome of the remaining 45%. Instead, I noted that reliable gains between time

points was a feature across therapy: between 10 and 20 sessions (20% improved); between 20 and 30 sessions (17.1% improved); and between 30 and 40 sessions (20.8% improved). Indeed, aggregated scores of 39 participants who had completed 40 therapy sessions demonstrated not only reliable improvement in the first ten sessions of therapy but also, following a plateau (10-20 sessions) and a period of nonreliable change (20-30 sessions), a second reliable gain during the final ten sessions of therapy. This pattern of change does not support the GEL model in which quickly improving participants leave the pool as they reach their GEL, thereby slowing down the aggregated estimates of the effectiveness of later sessions (Barkham et al., 2006), as all participants remained in therapy. Instead, it indicates a strong trend amongst participants for improvement in congruent functioning in both the early and late stages of therapy and either little change or, conversely, varied and contradictory trajectories amongst participants during the middle stages. This pattern of change is very similar to Owen et al.'s (2015, p.822) class 1 'early and late change' trajectory, the most represented of the three models that they identified in their study.

The similarity of this finding was supported and further developed by the identification of a cubic polynomial model as the best fitting non-linear model for the data in this study. The results confirmed that the change in participants' SI scores during the course of therapy demonstrated a high degree of variation: not only different rates of change but also different shapes, consistent with the multiple trajectories of change proposed by Owen et al. (2015).

Finally, the results of this study have provided preliminary evidence that scores on the Strathclyde Inventory do appear to change over the course of therapy in a way that is consistent with person-centred theory. These are the main findings arising from this study and will be discussed in greater detail in Chapter 7.

This study has confirmed that the SI is measuring something distinctively different to change in the individualised symptoms of distress captured by the PQ and the risk to self and others recorded by the risk scale of the CORE-OM. There is a clear association with the non-risk scale of the CORE-OM, which seeks to measure,

and cannot separate, an interlinked experience of distress and functioning. Person-centred theory of change, while not prioritizing the reduction of distress, clearly acknowledges a relationship between increased congruent functioning and decreased distress. Rogers (1951, p.513) wrote: "The feeling of reduction of inner tension is something that clients experience as they make progress in 'being the real me' or in developing a 'new feeling about myself'". More recently, Warner (2017, p.109) agreed, stating "such processing and self-cohesion allow the development of a 'congruent' version of self that resonates with the person's whole-body experience and minimizes psychological symptoms." The results of this study have provided preliminary evidence that confirms this view, having identified an association between greater change on SI and lower post-therapy distress and risk measured by CORE-OM but further research is required to investigate the nature of any causal relationship.

The identification of two SI items that appeared to change in functioning – or meaning, from the perspective of participants – over the course of therapy indicated that changes occurred in participants' perception of the item, or understanding of themselves in relation to the item, by the end of therapy. Responding to these two items became harder for participants. In addition, a small negative pre-post effect size found for one item identified that some participants reported feeling less true to themselves by the end of therapy. This suggests a more realistic, if painful, appraisal of themselves within their current life situation, indicative of greater congruent functioning as discussed below.

I began this study with a hypothesis that I might find the presence of early deterioration in scores for some participants; paradoxical evidence of increasing congruent functioning as awareness of self, and ability to articulate self within the current situation (Purton, 2013), begins to grow. Despite an overall linear trend of gradual improvement, and strong evidence of reliable gains occurring for some participants early in therapy, I did find evidence of reliable deterioration in scores between data collection points on fifteen occasions. Two-thirds of these reliable deteriorations in scores took place in the first twenty sessions of therapy. As with

participants who experienced sudden gains, many of these participants (60%) ended counselling at that point so we do not know what might have occurred if they had continued. However, having looked at the experience of the seven participants that did continue in therapy beyond a reliable deterioration in score, this study has demonstrated that this decrease in score does not necessarily lead to deterioration overall. Further exploration of the experience of participants whose SI scores indicate this phenomenon during therapy is required in order to understand what was happening for those individuals and to discover if their experiences can be understood from a theoretical perspective.

Indeed, many researchers (e.g. Blanton & Jaccard, 2006; Kazdin, 2006) have recommended caution about what can be inferred from the 'arbitrary metrics' provided by measures. It is clear that my interpretation of data collected in this form may not in fact represent a change that was experienced by these participants in 'real life' or, indeed, if such a change was perceived by participants as beneficial or meaningful (Kazdin, 2006, p.48). Therefore, my final study sets out to investigate whether pre-post change recorded in scores on the SI is meaningful – and nonarbitrary – for a small selected group of participants from this dataset.

Chapter 6: A Meta-synthesis of a Systematic Case Study Series Examining Client Improvement and Deterioration in Therapy Identified by the Strathclyde Inventory

Chapter Overview

In the previous chapter I demonstrated that change in scores on the Strathclyde Inventory (SI) over the course of therapy followed a variety of different patterns. Based on the results of this study, it was clear that the SI was sensitive to change when used to collect data from a UK clinical population, and also that there was preliminary evidence that this change in SI scores occurred in ways that could be interpreted as consistent with person-centred theory as it relates to the process of developing congruent functioning. However, an important question remained: does the change in scores recorded on the SI accurately reflect a participant's therapeutic outcome? This is an essential concern for the validity of any outcome measure.

I decided to investigate this question through a series of systematic case studies, focusing on participants whose SI scores either reliably increased or decreased between the beginning and end of therapy. In addition, this process enabled me to identify any similarities and differences in their experiences that might explain their therapeutic outcomes from the perspective of congruent functioning, thereby confirming or developing person-centred theory. I anticipated that this study would provide me with the final ingredients required to answer my two overarching research questions: (1) are scores on the Strathclyde Inventory a valid measure of congruent functioning when used with a UK-based clinical population? and (2) do scores on the Strathclyde Inventory change over the course of therapy in a way that is consistent with person-centred theory?

Research Questions

This third study was designed as a two-stage process aiming to answer the first two questions below, with two supplementary questions intended to guide discussion:

1. Does the pre-post change in scores recorded on the SI accurately reflect the client's overall therapeutic outcome when assessed using the Hermeneutic Single Case Efficacy Design (HSCED) method?

2. What features of the therapeutic experience may explain client improvement or deterioration at the end of therapy as measured by the SI?

3. To what extent does the meta-synthesis of the HSCED results validate change in SI scores?

4. To what extent are therapeutic processes associated with person-centred theory found in the HSCEDs causally related to change in SI scores?

Study design. In the first stage of this study, I supervised a group of eight MSc students who each conducted a case study of one client selected from the Study 2 sample using the Hermeneutic Single Case Efficacy Design method (HSCED; Elliott, 2014, 2015). The eight clients were selected according to a set of inclusion criteria (see below) with the intention of forming two groups: four clients whose scores on the SI *increased* during therapy ('improvers') and four clients whose scores on the SI *decreased* during therapy ('deteriorators'). The eight dissertations produced as a result of this first stage of the study are listed in Appendix F.

In the second stage of Study 3, I carried out a meta-synthesis of this series of systematic case studies in order to identify patterns of similarity or difference in the features of the therapeutic experience for clients within each group and between the two groups.

Ethical Issues. As McLeod (2010, p.54) noted, case studies pose a higher degree of potential risk to participants than other forms of research because the nature of a case study requires that the researcher focuses on the details of an individual: their circumstances, relationships, difficulties, responses and reactions, and so on. McLeod (2010) recommended strategies and guidelines for the ethical conduct of counselling and psychotherapy case study research based on three core principles: obtaining informed consent from clients, maintaining confidentiality, and

avoiding harm to case study participants (both client and therapist). This guidance has informed the decision-making of all researchers who contributed to this study.

At the Strathclyde Counselling & Psychotherapy Research Clinic ('the research clinic'), obtaining consent from clients is a rolling process in which clients are invited to complete a general consent form as well as a detailed Release of Recordings consent form (Appendices G & H) before therapy begins and then after every tenth counselling session, at the end of therapy and at follow up. As the research clinic is building an archive of data that can be used in a range of potential future research, the consent forms are designed so that clients can control the ways in which their data will be used. They are invited to review (i.e. maintain, increase, decrease or withdraw) the range of ways that they permit their data to be used on each occasion that they meet with their researcher. A specific clause relating to case study research was introduced into the Release of Recordings consent form in 2011. Together with my first supervisor, I assessed that all clients selected for this study had provided sufficient consent to be included. This was part of the selection process (see inclusion criteria below).

Stage 1: The HSCED Group Project

Method

HSCED methodology. The Hermeneutic Single Case Efficacy Design method (HSCED; Elliott, 2014, 2015) is a form of systematic case study that adopts a structured, critical-reflective approach to examine and evaluate the outcome of therapy within an individual case with the aim of addressing three fundamental questions in psychotherapy research (Elliott, 2015):

- did the client change? (outcome research);
- was the therapy generally responsible for the change (efficacy research); and
- what specific factors (within or outside therapy) are responsible for the change? (change process research).

As Elliott (2014) described, the HSCED was designed as a practical reasoning system (not unlike those used in other professions such as law and medicine) for "testing causal inferences in naturalistic situations" (p.352). It is hermeneutic because it seeks to interpret a rich range of quantitative and qualitative data drawn from a variety of sources and perspectives (presented in a *rich case record*) about a client's experience in therapy in order to draw inferences that can lead to a plausible understanding of the client's outcome in therapy and the processes that brought it about (p.352).

Critical reflection is facilitated by a dialectical process of scrutinizing the rich case record for evidence that can be used to develop two opposing accounts of what occurred: first, the *affirmative case*, which seeks to demonstrate, using at least two types of direct evidence, that change took place and was causally influenced by the therapy; and second, the *sceptic case*, which presents alternative explanations for any apparent client change. Elliott (2014, 2015) proposed five lines of direct evidence that can be drawn from a comprehensive rich case record to form a plausible affirmative case (retrospective attribution, process-outcome mapping, within-therapy process-outcome correlation, early change in stable problems, and event-shift sequences) and eight lines of indirect evidence that may be used to develop a credible sceptic case (trivial or negative change, statistical artifacts, relational artifacts, expectancy artifacts, self-correction processes, extra-therapy events, psychobiological causes, and reactive effects of research).

Several concepts and methods drawn from the traditions of Western legal systems have shaped the methodology underpinning the HSCED method in particular concerning the testing of evidence and issues of proof (Stephen & Elliott, 2011). Therefore, the process of developing two opposing cases mirrors the adversarial approach traditionally adopted in Scottish, English and American legal systems, amongst others. Interrogating the evidence from two opposite positions permits the plausibility of each position to be evaluated. Allowing *cross-examination* of each case – in HSCED, after presenting its own *brief*, each case develops a *rebuttal* of the other's arguments – enables the strengths and

weaknesses to be tested. The affirmative case carries the *burden of proof*. In other words, as for the Pursuer in a civil case or the Prosecutor Fiscal or Lord Advocate in a criminal trial within the Scottish legal system, it is the side that makes the claim who carries the responsibility of proving their claim to the *standard of proof* required: in HSCED this burden is carried by the affirmative case, which claims are that change has taken place and that this was a result of the client participating in therapy. The responsibility of the opposing side (in HSCED, the sceptic case) is to undermine their case sufficiently for this standard not to be met. In civil cases in Scotland, the standard of proof is *on the balance of probabilities* (i.e. the probability is greater than 50%) whereas in criminal cases the standard of proof is much higher: *beyond all reasonable doubt*. Elliott (2014, p.353) argued that this standard equates to near certainty and can be understood as the equivalent of the $p<.05$ standard applied in statistical significance testing. In contrast, Elliott (2014) proposed that an intermediate standard of proof should be customarily applied in HSCED: *clear and convincing evidence* (Stephen & Elliott, 2011, p.238; see also Chapter 5) that is, the probability that the claim has been proved is greater than 80% ($p<.2$).

The final characteristic of the HSCED method that reflects its quasi-judicial heritage is its use of an adjudication process in order to make a decision on the two cases presented. This aspect of the method has grown and developed since it was first introduced. HSCED was originally conceived as an informal critical-reflection method that could be applied by a therapist when reflecting on their work with a client, providing an activity that would help to bridge the practice-research gap (Elliott, 2014). The individual practitioner would make their own decision on whether their case had been proved. It can still be used in this way. However, one of the learnings from applying the HSCED method is that the question of whether or not the client improved is in fact highly complex (Elliott, 2014, p.358). For this reason, it is now generally accepted that an HSCED case seeking to be taken seriously should be considered by an adjudication panel (Benelli, De Carlo, Biffi & McLeod, 2015). The constitution of the panel has been experimented with over time: for example, distinguished psychotherapy researchers (Elliott et al. (2009);

experienced practitioner-researchers (Stephen, Elliott & MacLeod, 2011); trainee counselling psychologists (MacLeod & Elliott, 2014); and by including *experts by experience*, people who have personal experience of using or caring for someone who has used mental health services (Traynor, 2019).

Development of the panel's decision-making process has also been attempted. When first introduced, panel members considered the cases provided by the HSCED researcher independently and returned a decision based on their own individual assessment of the coherence of the arguments. It would then be the responsibility of the researcher to choose a method to consolidate the three or more decisions received. To assist members of the adjudication panel and to introduce a framework that would assist with both the standardisation and elaboration of the adjudication process, I developed a pro forma that invited panel members to use a rating scale to record their assessment of the evidence on each question – enabling them to go beyond a "yes/no" response to whether the affirmative case had met the 80% standard of proof on the question - and to outline the reasons for their decisions (Stephen et al., 2011). An example of the HSCED Decision pro forma used in this study, which is an adaptation of my original form, is included as Appendix I. The use of this form enabled me to more easily compare and contrast the three separate adjudications. The rating of the *extent* of client change, and the influence of the therapy on this, provided an opportunity to identify an *average* of the three panel members' decisions, in the case the median score on each rating scale, in order to represent the panel as a whole. However there remains further opportunity to develop the adjudication process to find a way for the panel to make a joint decision, closer to the consensus theory of truth (e.g. Barker et al., 2016, p.10), and recommended by Benelli et al. (2015). In the current study, the members of each adjudication panel were asked to work together to achieve a *decision by consensus* on each question.

There was one other major change to the method tested in this study. Since the first HSCED cases, the affirmative case was focused on demonstrating that a particular kind of change had taken place: improvement. Most published HSCED

cases have analysed *good outcome* cases (e.g. Macleod & Elliott, 2014) or *mixed outcome* cases (e.g. Stephen et al., 2011). Only one *poor outcome* case is known to have been published (MacLeod & Elliott, 2012). In this case the researcher applied the usual approach, that is for the affirmative case to prove that the client had improved. In consultation with my first supervisor, I decided in this study to 'flip' the method for the cases in which, based on their SI scores, the client had deteriorated. Therefore, the researchers conducting the HSCEDs for the four clients whose pre-post scores on the Strathclyde Inventory deteriorated set out in their affirmative cases to argue the *prima facie* case that the client had deteriorated, rather than improved, during therapy.

Participants

Clients. Eight clients (four improvers and four deteriorators) were selected from the Study 2 sample using the following inclusion criteria:

1. Reliable change (improvement or deterioration, $p < .2$) as defined by Jacobson and Truax (1991)
2. Standardised residual gain (predicted post-test score compared with actual post-test score)
3. Client gave consent for researchers to analyse and present their data
4. Number of sessions (diversity if possible)
5. Had ended therapy
6. Gender (diversity if possible)
7. Age (diversity if possible)
8. Recording of end of therapy (or final) change interview
9. Sufficiently rich descriptions in Helpful Aspects of Therapy forms
10. Recordings of therapy sessions

Initially, the group of four improvers selected from the sample (N=105; Table 5.9) were the four clients with the best standardised residual gain scores who met these criteria in full. The sample of deteriorators was significantly smaller (N=7). Three of these clients were excluded as they had not consented to their data being used for a project of this kind. I decided to include all four of the remaining deteriorators in

the study, despite a variety of missing data, in order to achieve a balance in numbers with the improvers group. However, I noticed that all four deteriorators had worked with more than one therapist; therefore, in order to balance the groups somewhat and increase the potential for exploring the impact of changing therapists, I returned to the improver group and replaced the third and fourth selections with the next two improvers who met the inclusion criteria and had also worked with more than one therapist.

This final selection of clients appeared relatively well balanced with each group including 2 males and 2 females and representing a range in the duration of therapy (improvers' mean sessions = 29.25; range 9-47; deteriorators' mean session = 35.5; range 16-55). All clients were White-European. Table 6.1 presents SI outcome data (standardised residual; reliable change) and demographic data for the selected clients, presented in order of their standardised residual scores.

Table 6.1. The eight clients selected for the HSCED project

Client	Strathclyde Inventory		Number of sessions	Gender	Age	Dissertation author (all 2018)
	Standardised residual	Reliable change				
Julia	2.11	2.33 (+)	9	F	23	Khan
Linda	2.02	1.97 (+)	20	F	50	Bell
James	1.74	2.13 (+)	41	M	46	Whitehead
Simon	1.68	1.75 (+)	47	M	47	Martin
Sofia	-1.21	.88 (-)	20	F	30	Price
Joseph	-1.36	.67 (-)	16	M	31	Mackintosh
Luke	-2.12	.87 (-)	55	M	18	Love
Caitlin	-3.09	1.74 (-)	51	F	19	Moran

Notes. (+) = reliable improvement; (-) = reliable deterioration.

Five of these clients were identified as outliers in Chapter 5: Julia (participant 665), Linda (participant 32), Luke (participant 316) and Caitlin (participant 300) were identified as having particularly large standardised residual scores (Table 5.11); Sofia (participant 416) recorded a pre-therapy SI score that was significantly higher than the majority of participants (Figure 5.2). In addition, James (participant 471) and Caitlin (participant 300) were amongst the group of seven participants who

continued in therapy after their SI scores had reliably deteriorated during therapy (Table 5.15).

Therapists and researchers. Twenty-six volunteers in the Research Clinic worked as either therapist or researcher with one or more clients in the sample. Six of the eight clients worked with more than one therapist: five clients had one change of therapist during their period of therapy with the Research Clinic; one client (Simon) experienced two changes of therapist.

It was also typical in this sample for the researchers assigned to the clients to change during their period in therapy: four clients had a change of researcher while in therapy; two of these clients changed researcher twice. One volunteer worked as a researcher with one client (Simon) but transferred to become their therapist when their second therapist left the service.

The majority of volunteers (n=23; 88.5%) were in training as person-centred counsellors or as counselling psychologists learning person-centred therapy in the first year of their doctorate course; the remaining volunteers were experienced therapists. One volunteer worked with two clients in the sub-sample: one when in training; one post-training. They were predominantly female (n=23; 88.5%) and had a mean age of 35.2 years (range: 23-58 years). There was some ethnic diversity among the volunteers: Asian-Chinese (n=2; 7.7%); Asian-Pakistani (n=1; 3.8%); Middle Eastern (n=2; 7.7%). The remaining volunteers were either White-European (n=19; 73%) or White-Non-European (n=2; 7.7%).

HSCED researchers. Eight students volunteered to conduct HSCED studies in this project as part of a 60 credit Counselling Research Dissertation class offered within an MSc in Counselling and Psychotherapy at the University of Strathclyde. All of the students were female and had a mean age of 32 years (range: 23-61 years). One student was Asian-Pakistani (12.5%); two were White-Non-European (25%) and the remaining five were White-European (62.5%).

All of the students were in practice as trainee person-centred therapists as part of their MSc course. Two of the group were therapists and researchers at the

research clinic so were familiar with the type of data collected from clients. Neither of them had been in practice at the research clinic when the eight clients were participating in their therapy.

As person-centred therapists in training, the HSCED researchers entered the study with an awareness of their potential bias in favour of finding that the therapy had been effective for the individual clients. Before beginning the process, I asked the HSCED researchers to reflect on their existing assumptions in relation to the 'authority' of quantitative versus qualitative data and what 'substantial change' (whether improvement or deterioration) would look like.

Supervisor. I designed and supervised the study for the purpose of creating data that I could use within my PhD investigation. Therefore, I was invested in the study and in supporting the HSCED researchers to understand and apply the adapted HSCED method to the best of their abilities within the overall pressure of their course. I had completed my own HSCED study using data from the research clinic to obtain an MSc in Counselling in 2008-10 (Stephen et al., 2011). Thus, I had personal experience of the challenges involved, but these students were working within a much shorter and more pressurised timescale than I had. I met with the HSCED researchers as a tutorial group on eleven occasions between March and July 2018. During these meetings we discussed the application of the research method to their individual cases and explored questions as they arose through the process. I was also available for consultation by email if required.

As a group supervisor within the research clinic I was able to recall varied degrees of detail about the therapeutic experience of some of the clients selected for the study (James, Simon, Sofia). I had met with one of the clients (Joseph) as researcher for his Mid-1 change interview and another (Caitlin) as a researcher when she returned for further therapy at the research clinic. I had supervised the therapists of five of the clients (Julia, James, Simon, Sofia, Joseph) and therefore had my own impressions of their therapeutic work. As a result, I was aware of having my own assumptions about the clients' experiences in therapy and potential reasons for their outcome, which I sought to 'bracket' when working with the

HSCED researchers so that their interpretations of the data was not influenced by my experience of the clients or therapists.

Measures

Quantitative outcome data*.* Three instruments were used to capture quantitative outcome data: the SI (Freire, 2007), CORE-OM (Evans et al., 2002; Connell et al., 2007), designed to measure general concepts of distress and functioning, and the Personal Questionnaire (PQ; Elliott et al., 2016), an individualised instrument created by the client at the intake interview to itemise specific difficulties that they wished to address in therapy. All three instruments were administered by the client's researcher at regular intervals: the intake interview, mid-therapy change interviews (see below), the end of therapy change interview and at a change interview held six months after the end of therapy, if the client agreed to take part in the follow up process. In addition, the PQ was used as a weekly outcome measure, completed by the client at the beginning of every counselling session.

Qualitative outcome data. The clients completed Helpful Aspects of Therapy forms (HAT; Llewelyn, 1988) at the end of each counselling session, in which they were invited to write brief descriptions of specific experiences within the session that they found helpful or hindering. After every tenth session of counselling, clients participated in a change interview (Elliott, Slatick & Urman, 2001) with their researcher, who used a semi-structured interview schedule to find out about the client's experience of counselling so far, including changes they had noticed in themselves, their understanding about what had caused these changes, the personal strengths and limitations that positively or negatively affected their ability to use therapy, specific and general examples of helpful and hindering experiences in therapy as well as the impact of the research protocol. Where possible, the HSCED researchers were able to listen to and transcribe recordings of the change interviews. In most cases, when recordings were not available, the HSCED researchers were able to access brief notes made by the client's researcher. Finally, therapists completed a therapist session form at the end of each counselling

session. This form contained space to record the main events in the session including any unusual within-therapy or extra-therapeutic events.

Relational data. Relational assessments were carried out at the end of sessions 3, 5, and then every fifth session until the end of therapy. The client completed the Working Alliance Inventory (Hatcher & Gillaspy, 2006), which seeks to measure the emotional bond between client and therapist as well as agreement on the goals and tasks of therapy. In addition, client and therapist completed parallel versions of the Therapeutic Relationship Scale (Sanders & Freire, 2008; Carrick & Elliott, 2013), designed to measure the quality of the relationship from a person-centred perspective. The earliest client in the HSCED study had also completed the Relational Depth Inventory (Wiggins, Elliott & Cooper, 2012) and data collected on this instrument is included in their HSCED material.

Procedure

Selecting cases. Each HSCED researcher selected the case that they would study at our first tutorial group meeting. First, I presented them with the personal questionnaire items for the eight clients, without identifying if they were from the improver or deteriorator group. Each HSCED researcher indicated the client(s) that they were interested in working with based on these self-generated descriptions of their main difficulties at the start of therapy. If only one HSCED researcher was interested in a particular client then they were assigned to them. Next, I provided more information about the demographic details of the clients, the number of therapy sessions that they received and their reliable change and standardised residual scores on the Strathclyde Inventory, which revealed if they were improvers or deteriorators. This additional information assisted the remaining HSCED researchers to negotiate with their peers to select the client whose therapy they wanted to study. All HSCED researchers expressed satisfaction with the outcome of this process. The final decision for the HSCED researchers at this stage was to choose a pseudonym for their client. In most cases this choice was influenced by the demographic details that they had been given about their clients (e.g. age, ethnicity).

Consent. Having completed the selection process, I gave the HSCED researchers access to the consent form completed by their clients when participating in the research process so that they could satisfy themselves with the detail of the consent given for the way in which their data could be used. They also checked what instructions their client had given concerning the anonymisation of personal details (e.g. names of friends and family, places, occupations) that may be included in their data.

Creating the rich case record. The HSCED researchers accessed the quantitative and qualitative data collected during their assigned client's therapy at the research clinic as described above. In some deteriorator cases, to make up for missing data, I provided HSCED researchers with anonymised copies of email correspondence between client and therapist to better understand the context in which therapy was occurring. Each HSCED researcher made decisions about which data to include in their rich case record and how to present it, although they also had access to previous HSCED studies that they could consult and use as models.

Preparing the HSCED analysis. Once the rich case record was created, the HSCED researchers began to analyse the data it contained from two alternative perspectives in order to compile affirmative and sceptic cases that addressed the three HSCED research questions:

1. Did the client change (improve/deteriorate) substantially over the course of therapy?

2. Was this change (improvement/deterioration) substantially due to the effect of the therapy?

3. What factors (including mediator and moderator variables) might have been responsible for the change?

Affirmative and sceptic cases were each made up of three parts: brief, rebuttal (to arguments made in the alternative case), and summary narrative. The HSCED researchers working with the improver cases followed the traditional lines of direct evidence (affirmative brief) and the non-therapy processes that may account for

change (sceptic brief) proposed by Elliott (2014). The HSCED researchers working with the deteriorator cases followed the same procedures as far as possible but adapted them to argue that the change that took place was deterioration. It became clear during the analysis that a further line of evidence was required in order to take into account that, while some of these clients did not explicitly identify certain aspects of their therapy as unhelpful, the evidence in the data when considered from a theoretical perspective suggested that it could be related to deterioration in therapy. Following discussion within the deteriorator sub-group, it was agreed that this new line of evidence would be called: 'Examination of the Rich Case Record revealed aspects of therapy which could be considered hindering, even if the client did not demonstrate awareness of these.'

When discussing specific details of cases during the process of preparing their analyses, the HSCED researchers worked in two sub-groups – improvers and deteriorators - so that those from the other sub-group, who would form the adjudication panel, would not be influenced by early access to details of the case.

Conducting the HSCED adjudication. The adjudications were scheduled to take place in parallel sessions during the second and third last meetings of the tutorial group. HSCED materials were provided to members of the relevant adjudication panel in advance (range: 1 – 7 days). The first parallel session, in which one improver case was adjudicated by a panel of three students who had carried out deteriorator HSCEDS, and one deteriorator case was adjudicated by three students who had carried out improver HSCEDS, was treated as a pilot. The process took longer than the allocated time, providing an opportunity for the adjudicators to consider what changes they could make to their preparation and application of the process in order to work more efficiently with subsequent adjudication processes.

The adjudication panels were provided with a summary document called the HSCED Adjudication Panel Decision Pro Forma (2018 Version) on which to record their decisions and a brief narrative of the reasons for their decisions. A copy of this form has been included in the appendices (Appendix I). The questions were

expanded versions of the three HSCED research questions designed to help the adjudication panels approach each question from more than one perspective: (a) the *degree* of change that they judged the client made and (b) their *confidence* that the client had changed *at least substantially*. The panels were asked to indicate their answers using a rating scale that offered two alternative sets of labels to assess their response: language (no change, slightly, moderately, considerably, substantially, completely) and percentage (0-100% in 20% increments). The forms also provided opportunities for the panels to use narrative to note and explain the evidence that influenced their conclusions.

Results and Interim Discussion

The main decisions of the adjudication panels for each case study are presented in Table 6.2.

Table 6.2. Main decisions of the HSCED adjudication panels

		To what extent do you think the client changed (improved/ deteriorated) over the course of therapy?	To what extent do you think that the client's change (improvement/ deterioration) was due to the therapy?
Julia	↑	Substantially-Completely (90%)	Substantially (80%)
Linda	↑	Substantially (80%)	Considerably–Substantially (70%)
James	↑	Substantially (80%)	Considerably–Substantially (70%)
Simon	↑	Considerably-Substantially (75%)	Substantially (80%)
Sofia	↓	Slightly-Moderately (30%)	Slightly-Moderately (35%)
Joseph	↓	No change-Slightly (15%)	Moderately (40%)
Luke	↓	Moderately (40%)	Moderately (45%)
Caitlin	↓	Considerably (60%)	Moderately-Considerably (50%)

Notes. ↑ = 'improver'; ↓ = 'deteriorator'. Decision scale used by the adjudication panels: No change (0%); Slightly (20%); Moderately (40%); Considerably (60%); Substantially (80%); Completely (100%).

The case studies are presented in order of standardised residual scores as in Table 6.1. It is immediately noticeable on the first question that, with the exception of Sofia and Joseph, the HSCED decisions (represented by the language and percentage labels used by the HSCED decision pro forma) parallel this order.

The extent of change (improvement) for improver clients ranges from Substantially-Completely (90%) for Julia to Considerably-Substantially (75%) for Simon. This result reflected the standardised residual scores for these clients and therefore suggests that the SI has been successful in capturing the degree of change experienced by these clients over the course of therapy.

This is also true at the opposite end of the spectrum in the decisions made about the extent of deterioration in the cases of Luke (moderately; 40%) and Caitlin (considerably; 60%). It is interesting to note that the panels' assessment of the change experienced by these clients is not as 'extreme' as that indicated by their standardised residual scores (-2.12 and -3.09 respectively), especially when compared to the decision for Julia, the improver with the highest standardised residual (2.11). This raises three possibilities: first, that the full range of evidence of deterioration offered using the HSCED process was not as severe as that indicated by the change in SI score, providing a strong rationale for the Stage 2 investigation, as this suggested that perhaps the instrument captures only one aspect of the client's outcome in therapy; second, that the meaning of deterioration on the SI is different to the meaning of deterioration in therapy overall, or at least as applied by the panel; or, third, that the panels were more moderate in their judgements of the deteriorators' cases than the panels adjudicating the improvers' cases, perhaps an artefact of the design of the study: members of each panel came to the adjudication process having been emerged in preparing the opposite type of case. This last point was one identified and discussed by the HSCED researchers in their dissertations as a potential limitation of their investigations.

The results in the cases of Joseph and Sofia also raised questions. The standardised residuals were far smaller than the other two deteriorators (-1.35 and -1.20, respectively) but of note because both clients entered therapy with non-clinical scores on the SI (2.5 and 3.38, respectively) meaning that they experienced not only reliable deterioration but also their scores crossed into the clinical range by the end of therapy. The HSCED decisions suggest that the panels found it difficult to make a decision that deterioration took place in the case of Joseph, while the result

in Sofia's case seems to fit better into the general flow of decisions on the first question. This may be, as described above, an artefact of the SI, or it may reflect something in Joseph's case that influenced the adjudication panel. Again, I was encouraged that the next stage of the investigation would provide an opportunity to address these questions.

The panels' decisions on the second question – the extent to which the change was due to therapy – follows a similar, although slightly more moderate, pattern when compared to the first question, except in the cases of Simon and Joseph (bringing his result on this question in line within the group of deteriorators). This result implies that the adjudication panels for the improvers' cases were substantially convinced (70-80%) that the therapy were causally related to the change experienced by these clients, whose scores on the SI improved over the course of therapy. In contrast, the adjudication panels for the deteriorators' cases were moderately less convinced (35-50%) that the therapy had caused the deterioration experienced, perhaps reflecting the general tendency noted in attribution theory (e.g. Kelley, 1973) for humans to be more likely to see positive outcomes caused by internal factors (in this case, the therapy) and negative outcomes caused by external factors. The explanations offered by the adjudicational panels in these decisions, including their proposals of the therapeutic processes that may have facilitated the improvement or change experienced by the clients, along with the HSCED material on which they were based, have been analysed in the next stage of the investigation.

Stage 2: Meta-synthesis of the Case Study Series

Method

Once the eight HSCED studies were completed I was able to commence stage two of this study: the meta-synthesis. Barker et al. (2016, p.41) described a meta-synthesis as "a thematic analysis of thematic analyses", a qualitative version of meta-analysis, in which the themes arising from each source provides the raw data for an overarching analysis. More specifically, Iwakabe & Gazzola (2009, p.605)

recommended meta-synthesis as one of "three main avenues to aggregating and synthesizing case studies", along with computerised case database and systematic comparison, highlighting the main goals of meta-synthesis as theory building and development, and the systematic identification of shared concepts and themes across similar cases (p.606). When I had first envisaged this study, I had imagined using a systematic comparative approach, closer to the cross-case analysis applied by Widdowson (2013), but on close inspection of the HSCED materials generated by my co-researchers, I recognized that, despite the systematic nature of the HSCED method, the individual experiences of clients and perspectives of the HSCED researchers had resulted in differing foci and emphases. Therefore, I decided to conduct a meta-synthesis using a version of grounded theory analysis, inspired by Timulak (2007, 2009), which would allow overarching themes to cluster within, and emerge from, the data itself.

Data Collection. In this study, the raw data for the meta-synthesis were the products of the HSCED analyses: the affirmative and sceptic cases, and the adjudications. My rationale for this design was that the HSCED researchers had examined the 'raw' data collected from the client and therapist and had identified and assessed the main features that might have contributed to the client's process and outcome in the therapy. Their examination of the evidence was weighed and further interpreted by the adjudication panel, highlighting and summarizing the most plausible explanations. Therefore the data analysed within this meta-synthesis had already been through two layers of interpretation prior to my analysis with the advantage that it provided me with a condensed representation of the content of each case, saving me from beginning my analysis at the rich case record level, but also the disadvantage that there may have been significant differences in the way that the HSCED researchers and I selected and made sense of the raw data.

Even with the reduction of data available to me as a result of the HSCED process, I found that there was a great deal of data to analyse. The material produced by each HSCED researcher (affirmative and sceptic cases; brief, rebuttal and summary narrative) was substantial, ranging from 21 – 57 pages. I worked

through each case study extracting short sections of text that appeared to relate to my main exploratory research question: What features of the therapeutic experience might explain client improvement or deterioration at the end of therapy as measured by the Strathclyde Inventory? These 'meaning units' formed the building blocks of my analysis.

Data Analysis. First, I carried out a preliminary sorting process as a pragmatic means of bringing together similar meaning units from the full range of case studies into broad general categories (client, therapeutic relationship, impact on client, attribution of change, nature of ending) then sub-divided them into separate Word documents that made it more manageable for me to handle the meaning units and carry out the analysis. This framework developed as I worked through the case studies one by one, alternating between deteriorator and improver cases.

Next, I worked with the meaning units within each broad area comparing each meaning unit to each other meaning unit that I had extracted from the HSCED materials. I organized meaning units with the same or very similar meaning or focus together into clusters. This process of constant comparison (e.g. Corbin & Strauss, 2008) enabled me to recognize similarities and discern differences within the data and to create a hierarchy of categories and sub-categories, which I summarized in tables to assist me to step back, take in the bigger picture and notice and experiment with relationships between the groups of analyses that I had carried out (i.e. *axial coding*; Corbin & Strauss, 2008). As a result, I was able to reorganize my domains and categories into the version presented here.

Following Hill et al. (2005), the final stage of my analysis was to work through the analysis to present a frequency count that characterised the representation of each of the two groups, improvers and deteriorators, within each category and sub-category. This process allowed me to identify which aspects or features of the therapeutic experience were *general* for improvers or deteriorators (the category or sub-category contained meaning units representing all four case studies), *typical* (three of the four case studies represented), *variant* (two of four),

or *unique* (one of four). These frequency labels enabled me to compare the features of the therapeutic experiences for the clients in the two groups, in particular, highlighting features that were general or typical for each group and also identifying differences between the groups, as recommended by Hill et al. (2005), when the results diverged by at least two frequency categories (e.g. improvers – general; deteriorators - variant).

Results of Meta-synthesis

The results are organized into three domains based on points in time within the therapeutic experience: *In the Beginning*, *In the Process*, and *In the End*. I present the findings in this order, highlighting the similar and different features of the experience for improvers and deteriorators, as depicted in the HSCED materials. Although this has been a process of bringing together the perspectives of a variety of participants (clients, therapists, HSCED researchers and adjudicators), to simplify the writing and reading process, I have presented the findings in a simple narrative as if it represents the actual experiences of the clients who were the subjects of the eight case studies. I have provided examples of meaning units from each category and sub-category. Each meaning unit is labelled with three pieces of information: (1) the perspective of the 'speaker' (i.e. client, therapist, researcher or adjudication panel); (2) the name of the client who was the subject of the case study; and (3) the relevant part of the HSCED material from which it was extracted (e.g. affirmative brief; sceptic rebuttal; adjudication). At the end of each domain I present a short narrative summary of the findings within that section. An overview of the analysis with the frequency count for each category and sub-category for the whole group and the two sub-groups, improvers and deteriorators, can be accessed at Appendix J. The full analysis, presenting categories and sub-categories with the supporting meaning units, can be accessed at Appendix K.

In addition, I have prepared figures to depict the general and typical features of the therapeutic experience for each group, highlighting the similarities and differences. Figure 6.1 represents the experience of the group of improvers and Figure 6.2 the experience of the group of deteriorators.

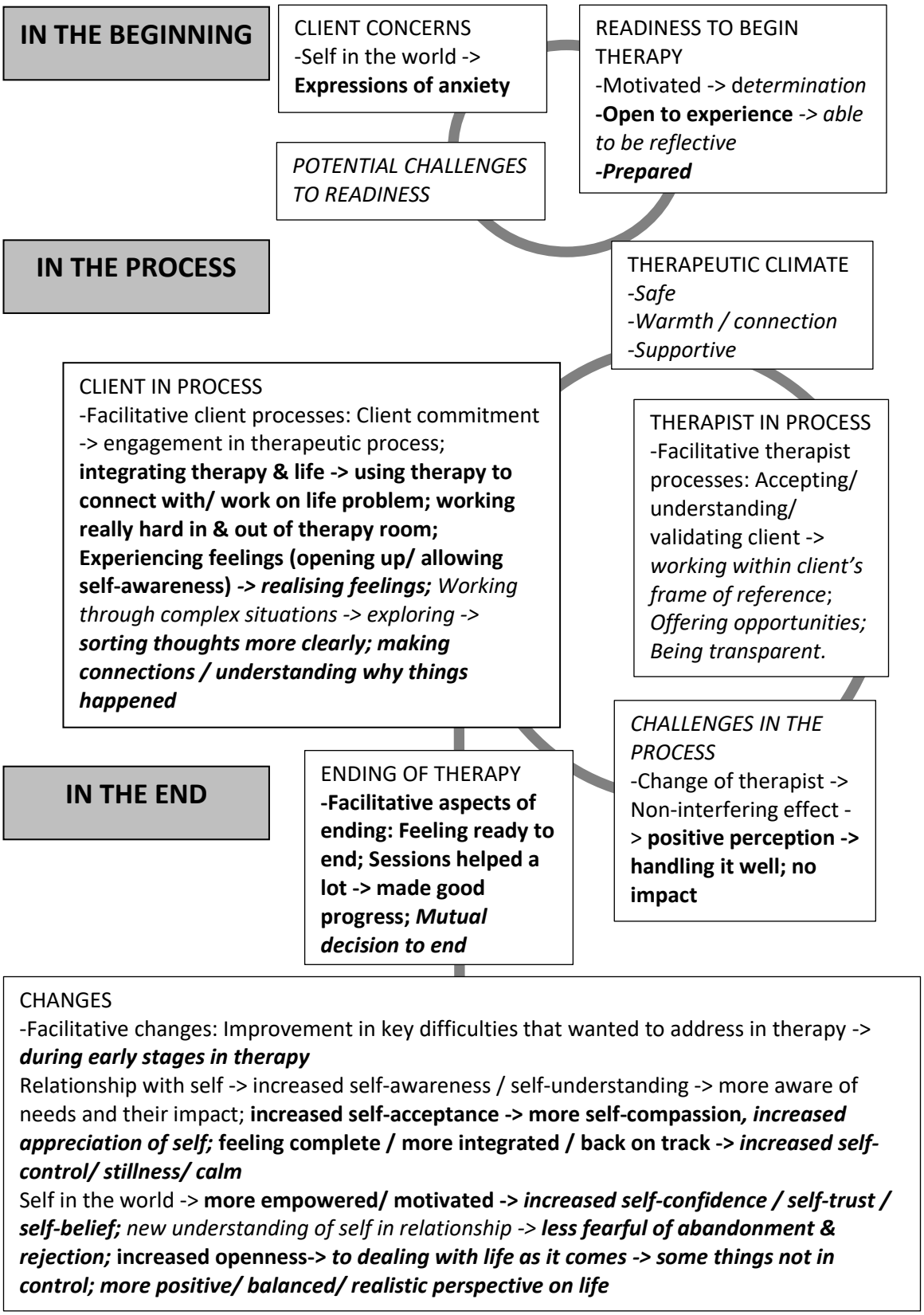


Figure 6.1 Improver Clients

Notes. **Bold** text = difference of at least two frequency counts; non-italicized text = general frequency; italicized text = typical frequency.

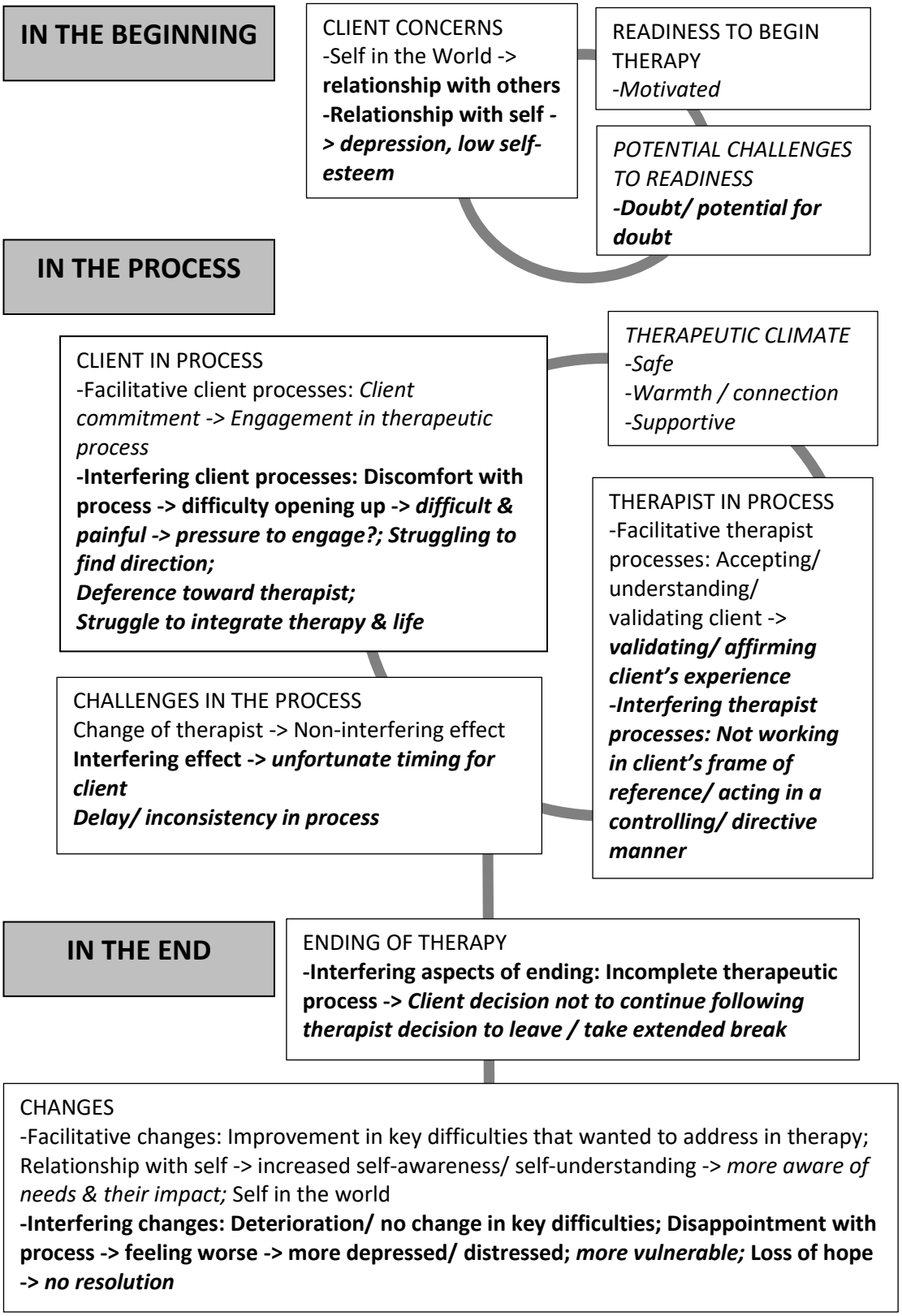


Figure 6.2. Deteriorator Clients
Notes. As for Figure 6.1

The three domains are presented in order from the top to the bottom of the figure. The main categories for each domain are shown and linked by a connecting line. The general and typical findings within each category are presented: general findings are non-italicised; typical findings are italicised. Those findings that are different to the findings for the other group by at least two frequency counts are presented in bold.

Domain 1: In the Beginning

The first domain describes features of the clients' experience as they came into therapy and is organized into three categories: *client concerns*, *readiness to begin therapy* and *potential challenges to readiness*.

Category 1.1: Client Concerns. This category contains the main concerns that brought the clients into counselling. These might have been recorded in a variety of ways: for example, as items in the client's Personal Questionnaire, disclosed in change interviews or referred to in therapist session notes or client post-session HAT forms. I found that these *client concerns* could be sorted into two sub-categories: *self in the world* and *relationship with self*. I was interested to note that this way of organizing client concerns highlighted immediate differences between the groups of improvers and deteriorators. Although both groups were fully represented in the sub-category, *self in the world*, the focus of their concerns were different: improvers' difficulties were experienced as *expressions of anxiety* (improvers – general; deteriorators – variant), whereas deteriorators perceived their difficulties in terms of their *relationship with others* (improvers – variant; deteriorators – general).

Sub-category 1.1.1: Self in the World. All eight clients experienced concerns about *self in the world* (improvers – general; deteriorators - general). These concerns were presented in the form of *expressions of anxiety* (improvers – general; deteriorators – variant), in *relationship with others* (improvers – variant; deteriorators – general) and as responses to *contextual difficulties* (improvers – unique; deteriorators – variant).

All improvers described a variety of experiences that could be understood as an *expression of anxiety*, for example:

> She reported experiencing long standing social anxiety difficulties which were hindering her daily life activities (Researcher; Julia, Affirmative Narrative)

> Yeah, I wasn't even aware actually that I was in a state of perpetual panic because I was so used to it. (Client; James, Affirmative Brief)

Six clients described difficulties that they had in their *relationships with others* in their life (improvers – variant; deteriorators – general).

> It became clear that there were significant emotional difficulties in his home. (Researcher; Luke, Sceptic Brief)

> Difficulties in all relationships – family/friends/partner (Adjudicating panel; Caitlin, Adjudication)

Some clients described the impact of these relationship difficulties in their PQ items and in therapy: *feelings of isolation* (improvers – unique; deteriorators – unique) and a *tendency to take on other people's issues* (improvers – unique):

> PQ items: 1) I feel I am invisible. 2) I have no sense of belonging. 4) I feel like an afterthought.6) I don't have anyone to fulfil my emotional needs. 9) I don't feel appreciated. 10) I feel lonely. (Client; Linda, Affirmative Brief)

Some clients identified that their concerns were linked to specific *contextual difficulties* that they were experiencing: *work-related problems* (improvers – unique; deteriorators - unique), and difficulty in *adapting to life in the UK* (deteriorators – unique). For example, in the case of Sofia:

> In Sofia's Intake Interview she identified the main problems that led her to seek therapy on the Problem Description Form. These included "existential anxiety around 'where is home?'" (Researcher; Sofia, Affirmative Brief)

> She had a number of issues surrounding her transition to life in the UK. (Researcher; Sofia, Affirmative Narrative)

Sub-category 1.1.2: Relationship with Self. Finally, six clients experienced issues connected to their relationship with themselves (improvers – variant; deteriorators – general). For deteriorators it was typical for this to be expressed as *depression* and *low self-esteem.* For example:

> Joseph reported difficulties with depression (Researcher; Joseph, Affirmative Narrative)

> Luke was in a very low place when he came to therapy, feeling very worthless and hopeless. (Researcher; Luke, Affirmative Narrative)

Category 1.2: Readiness to Begin Therapy. All clients, improvers and deteriorators, displayed at least one characteristic of *readiness to begin therapy*: my analysis suggested that they were *motivated* (improvers – general; deteriorators – typical), *open to experience* (improvers – general; deteriorators – variant), and *prepared* (improvers – typical; deteriorators – variant). There seemed to be a greater 'robustness' around the readiness experienced by improvers compared to deteriorators as it was typical for improvers to be represented in every sub-category, creating a web of readiness.

Sub-category 1.2.1: Motivated. Seven clients were clearly *motivated* to get something from the process (improvers – general; deteriorators – typical). This motivation was expressed in a number of ways: their *determination* (improvers – typical; deteriorators – variant), *expectations* (improvers – variant; deteriorators – variant), and *hope* (improvers – variant; deteriorators – unique).

> Throughout Linda's Change Interviews, she repeatedly indicated an expectation and a personal determination to change. This shows that she had significant expectations from the therapeutic process, and a strong sense of personal agency in making changes. (Researcher; Linda, Sceptic Brief)

> She identifies her own stubbornness and how this may influence her desire to change (Adjudicating panel; Julia, Adjudication)
>
> The client appeared to be determined to continue with the therapeutic process, showing agency in seeking appointments, and communicating extensively over email. (Adjudicating panel; Luke, Adjudication)
>
> When she came to therapy, she was struggling to cope with these difficulties but was hopeful that therapy could bring some form of resolution. (Researcher; Caitlin, Affirmative Narrative)

Sub-category 1.2.2: Open to Experience. There was evidence in the data that six clients (improvers – general; deteriorators – variant) were coming into therapy with the capacity to be open to their experience. This was characterized by being *able to be reflective* (improvers – typical; deteriorators – variant). For example:

> He was able to articulate his thoughts clearly and explain what his processes were. (Adjudicating panel; James, Adjudication)
>
> I think that I've got a certain way of thinking, whether it's the career or it's a natural reflection. What started it was an incredible shyness, introspection and then someone saying I was reserved. (Client; Simon, Sceptic Brief)

In addition, some of these clients demonstrated from the beginning of therapy that they were *willing to be open* (improvers – variant; deteriorators – unique) and *prepared to go to 'unhappy places'* (improvers – unique; deteriorators – unique):

> The extensive content of [her HAT forms] as well as other qualitative data obtained from Change Interviews […] support the notion that Caitlin was invested in therapy and willing to use sessions to be open and reflective. (Researcher; Caitlin, Affirmative Brief)

> No, I think you've got to go there... I think I was aware that with therapy, you have to go to the unhappy places to get resources. (Client; Linda, Sceptic Brief)

Sub-category 1.2.4: Prepared. There was more evidence of improvers being *prepared* to begin the process than deteriorators (improvers – typical; deteriorators – unique). This preparation was demonstrated in a variety of ways: these clients were *ready to get started* (improvers – variant), *proactive* (improvers – variant) or *knew what they wanted to work on* (improvers – unique; deteriorators – unique):

> I was ready for change when I came here. I knew I had to change. If I was to go on, then I had something... Something had to change. (Client; Linda, Sceptic Brief)

> I'm going to take copies of PQs away with me and make some notes for next session. (Client; Simon, Affirmative Brief)

> Knew what she wanted to work on. (Adjudicating panel; Sofia, Adjudication)

Category 1.3: Potential Challenges to Readiness. There was evidence of *potential challenges to this readiness* to begin therapy that was equally present in both groups (improvers – typical; deteriorators – typical). Only one of the sub-categories– *doubt/potential for doubt* (improvers – unique; deteriorators – typical) - was sufficiently supported by data to be identified as at least a typical aspect of the experience of one group and indicated a difference between the groups. The other sub-categories - *difficulties in processing emotions and experience, tendency toward rigidity* and *doubt, lack of supportive relationships* - offer examples of the type of challenges that may be around for any client when beginning the therapeutic process.

Sub-category 1.3.1: Doubt / potential for doubt. Some clients had *doubt* (improvers – unique; deteriorators - unique) - or had the potential for experiencing doubt because of an *expectation that would feel more vulnerable* (deteriorators –

unique) or *'wrong' expectations* (deteriorators – unique) - about whether the process would be useful for them (improvers – unique; deteriorators – general):

> So, I thought it was really just the job. It really was just the pressure. I didn't like where I was, so I felt much better. So, by the time I came to coming here, I wasn't going to come, because I thought I didn't need it now. (Client; Linda, Sceptic Brief)

> Caitlin had an expectation of increased feelings of vulnerability because of attending counselling (Researcher; Caitlin, Affirmative Brief)

> Yeah, well, I dunno, it's not that the therapy's unhelpful at all, it's great to have someone to talk to but then, maybe I was, maybe I had the wrong expectations but it was, from the beginning pretty clear that it's me doing all the talking. Sometimes it feels like, you kinda want that advice or that constructive feedback, a bit more honest. (Client; Joseph, Sceptic Brief)

Sub-category 1.3.2: Difficulties in processing emotions and experience.

There was evidence that four clients (improvers – variant; deteriorators – variant) came into therapy with some degree of difficulty in processing their emotions and experience. These difficulties took various forms: *accessing and expressing emotions / experience* (improvers – unique; deteriorators – variant), *understanding their own experience* (improvers – unique; deteriorators – unique), and *wanting to avoid difficult emotions* (improvers – unique; deteriorators – unique). For example:

> The boulder that was shutting my emotions - I had to get rid of the boulder. (Client; James, Affirmative Brief)

> He continued to find it difficult to symbolise his experiences. (Researcher; Joseph, Sceptic Brief)

> Luke was extremely detached from his own experience, writing "When I try to talk about such things it feels to me as if I was

> somebody else's story." This hindered him from bringing content into therapy; he wrote that he could "barely remember what exactly happens and it's even harder to explain to somebody else." (Researcher; Luke, Sceptic Rebuttal)
>
> I don't really want to talk too much about that because it tips my balance to painful negativity; just talked briefly about it. (Client; Simon, Affirmative Brief)

Sub-category 1.3.3: Tendency toward rigidity. Early in the process, two clients (improvers – unique; deteriorators – unique) shared parts of themselves that suggested a tendency toward rigidity that may have challenged them during their time in therapy: *perfectionism* (improvers – unique) and *unwillingness to consider/accept another's point of view* (deteriorators – unique). For example:

> The most frequently referred to hindering personal characteristic expressed by Caitlin was that she believes herself to be stubborn and unwilling to accept or consider another individual's point of view (Researcher; Caitlin, Sceptic Brief)

Sub-category 1.3.4: Lack of supportive relationships. Two deteriorators entered therapy with an obvious *lack of supportive relationships* in their lives.

> Luke's circumstances were hindering as well, notably his feelings of isolation. This was demonstrated by him spending his 20th birthday alone, and a growing awareness that he was going to fail his exams because he had not been attending lectures and had nobody to ask for notes (Researcher; Luke, Sceptic Rebuttal)

In the Beginning: Summary. In this summary section, and those for subsequent domains, my purpose is to highlight the *general* and *typical* findings within this domain for each group, and to identify the main differences between the groups (i.e. where the results for each group diverged by at least two frequency counts). These findings are also presented in Table 6.3 in which the main differences between the groups are highlighted in bold.

Table 6.3. Domain 1: In the Beginning

Category	Improvers	Deteriorators
Client concerns	Self in the world -> **Expressions of anxiety**	Self in the World -> **relationship with others** **Relationship with self -> *depression, low self-esteem***
Readiness to begin therapy	Motivated -> d*etermination* **Open to experience** *-> able to be reflective* ***Prepared***	*Motivated*
Potential challenges to readiness		***Doubt/ potential for doubt***

Notes. **Bold** text = difference of at least two frequency counts; non-italicized text = general frequency; italicized text = typical frequency.

Improvers entered therapy with concerns relating to *self in the world;* in particular, experienced as *expressions of anxiety*. Unlike deteriorators they demonstrated a robust "web" of *readiness to begin the process,* being *motivated, determined, open to the experience*, *able to be reflective,* and *prepared*.

Deteriorators also brought concerns about *self in the world*, but in contrast their concerns focused more on their *relationship with others.* In addition, they identified difficulties in their *relationship with self*, typically experienced as *depression* and *low self-esteem.* Like the improvers, most deteriorators were *motivated* to participate in therapy, indicating some degree of *readiness to begin the process.*

There was evidence that both groups experienced a variety of *potential challenges to their readiness* to begin therapy. The strongest finding was that deteriorators were more likely than improvers to enter therapy with *doubt / potential for doubt*.

Domain 2: In the Process

This domain describes the therapeutic process that developed between the clients and their therapists. It contains four categories: *therapeutic climate, therapist in process, client in process,* and *challenges in the process*.

Category 2.1: Therapeutic Climate. The evidence suggests that clients in both groups experienced the climate in which the therapy took place as *safe,* containing *warmth/connection*, *supportive* and, for some, *collaborative.* I found no distinct differences between the two groups in the aspects of the therapeutic climate described.

Sub-category 2.1.1: Safe. Clients commented on their perception of safety in the therapeutic relationship that they developed with their therapist (improvers – typical; deteriorators – typical). For example:

> And then there came a point when I thought, no, maybe there is [sic] issues that I really need to bring up, release, and let go in this safe environment. And really deal with things. (Client; Linda, Affirmative Brief)

> I felt that it was safe to share/express my feelings in a more emotional way and not just in an academic way. (Client; James, Affirmative Brief)

This sense of safety opened up possibilities for their participation in the therapeutic process in a variety of ways. They were able to*: express feelings without upsetting someone* (improvers – unique; deteriorators – unique); *be real/authentic* (improvers – unique; deteriorators – unique); *experience relational/emotional depth* (improvers – unique; deteriorators – unique); and *show vulnerability* (improvers – unique). For example:

> When asked to identify helpful aspects in the therapeutic process Joseph is able to offer several examples which relate predominantly to the opportunity for him to talk and share his experiences with a neutral person. (Researcher; Joseph, Sceptic Brief)

> It was a safe space so I could be me, exactly me, and not an act. (Client; Caitlin, Sceptic Brief)

Sub-category 2.1.2: Warmth/Connection. It was possible to infer the warmth and connection that developed between the clients and their therapists (improvers – typical; deteriorators – typical) from comments they wrote in post-session forms and, in the case of Luke who usually did not complete HAT forms, from email correspondence between them. For example:

> Since the last session I had become much more aware of my therapist as a person. (Client; James, Affirmative Brief)
>
> He said he wanted me to see it as I had mentioned before that there were parts of his life I didn't know about. I was quite touched that he made such a big effort. (Therapist; Simon, Affirmative Rebuttal)
>
> Luke's last emails to and from the second therapist, as well as the notes from the last session demonstrate a degree of warmth and prizing. (Researcher; Luke, Sceptic Rebuttal)

Sub-category 2.1.3: Supportive. A third quality of the therapeutic climate that became apparent when analyzing the data was its *supportive* nature (improvers – typical; deteriorators – typical). This was experienced both within sessions – in particular, *space to process* (improvers – unique; deteriorators – unique), *opportunity to practice communication skills* (improvers – variant) – and as *support between sessions* (deteriorators – variant). For example:

> The biggest help that we could see was the therapeutic relationship and the offering of a space where Simon could process what was important to him. (Adjudication panel; Simon, Adjudication)
>
> Additionally, Sofia has contact with the first therapist outwith the session; the therapist sends her articles and links to material she thinks may be relevant to Sofia and Sofia also contacts the therapist regarding a medical condition she is concerned about. (Researcher; Sofia, Affirmative Brief)

Sub-category 2.1.4: Collaborative. Collaboration between client and therapist was noted as a feature of the therapeutic relationship for a minority of clients in both groups (improvers – variant; deteriorators – variant). This could be seen in the various ways in which the clients or therapist described their work together, for example:

> She mentioned "we" on a couple of occasions which indicates the collaboration between client and therapist. (Adjudication panel; Julia, Adjudication)

> But that was good because it made me go, "what do you want me to talk about, what do you want me to say?" (Client; Joseph, Affirmative Rebuttal)

Category 2.2: Therapist in the Process. I organised the descriptions of the therapist within the therapeutic process into two sub-categories: *facilitative therapist processes* and *interfering therapist processes.* My identification of the process as facilitative or interfering was based on the way in which it was described by the original source of the meaning unit: client, therapist or researcher. All clients were represented in the sub-category *facilitative therapist processes*, whereas data from only four clients (improvers – unique; deteriorators – typical) were included in the sub-category *interfering therapist processes.*

Sub-category 2.2.1: Facilitative therapist processes. My analysis identified three types of *facilitative therapist processes*: *accepting / understanding / validating the client*, *offering opportunities* and *being transparent.*

Sub-category 2.2.1.1: Accepting / understanding / validating client. There was evidence that every client's therapist (or at least one of their therapists) was able to accept their client within the therapeutic process (improvers – general; deteriorators – general). For example:

> Therapist acceptance allowed James to reduce feelings of guilt and shame and he was able to have a rewarding interpersonal experience. (Researcher; James, Affirmative Narrative)

> She accepted him back, week after week – not making herself an enemy no matter what Luke did. He expected people to turn on him eventually, stating [in an email] "It seems that every person I've ever met became my enemy". It was probably incredibly healing for Luke to have someone accept him despite his anger towards them. (Researcher; Luke, Sceptic Brief)

My analysis suggests that this acceptance was conveyed in three particular ways. The first characteristic was *working within the client's frame of reference* (improvers – typical; deteriorators – variant). This was demonstrated through post-session notes written by clients and therapists about specific interactions between them, highlighting the ways in which the therapist stayed with the client's imagery or expressions, moment by moment, sometimes sharing their intention. For example:

> I reflected that there was one tree (according to client's image) that still needs to be knocked down, and added that I don't want to push the client to do that and hope I haven't today. (Therapist; Linda, Affirmative Brief)

> I think it was good of therapist to remind me that it's up to me to choose what I want to talk about in the sessions. (Client; Simon, Affirmative Brief)

There was evidence of the therapist expressing their understanding of the client's frame of reference through *using metaphor* (improvers – variant), *reflecting words back* (improvers – unique), and *reflecting feeling* (deteriorators – unique). For example:

> [What was helpful in the session?] The part about navigating anxiety being like an acrobat on a bicycle on a wavering line, where falling would be falling into anxiety. (Client; Julia, Affirmative Brief)

> Counsellor reflects how much anger is present for Caitlin. (Researcher; Caitlin, Sceptic Brief)

A second way in which therapists conveyed their acceptance of their client was through *validating/affirming client's experience* (improvers – unique; deteriorators – typical). This was the only sub-category within *facilitative therapist processes* in which the analysis found a difference of at least two frequency counts between the two groups. Clients described feeling that their therapist's response to them when describing their difficulties was reassuring as they confirmed that their experiences could be taken seriously and also, for some, that they could change. For example:

> I talked freely for some time and therapist agreed that there was a lot going on. This helped me to decide to look at my life and decide that some things are in the past. (Client; Simon, Affirmative Brief)
>
> It wasn't my intention to upset her but it validated that what I was saying out loud, that it mattered and it was upsetting. (Client; Joseph, Sceptic Brief)

For some clients, a third demonstration of the therapist's acceptance was their attention to *helping the client feel comfortable* in the process (improvers – unique; deteriorators - variant). For example:

> It was important that the therapist made me feel comfortable with being involved in the process. (Client; Simon, Affirmative Brief)
>
> In session 19, she also wrote about how comfortable she felt suggesting things in therapy: 'the fact that I suggested a new way of dealing with my concerns - comfortable enough to feel that this is something I can place as a proposal in the therapy space'. (Researcher; Sofia, Sceptic Brief)

Sub-category 2.2.1.2: Offering opportunities. There was evidence of the therapists responding to their clients in ways that seemed to be *offering opportunities* (improvers – typical; deteriorators – variant) within the process. There was no typical way of doing this. In my analysis I characterised these offers in four ways: *to ground/slow down the client* (improvers – variant; deteriorators – unique)*, to develop the process* (improvers – variant; deteriorators – unique)*, to engage the*

client (improvers – unique; deteriorators – unique), and *to challenge the client* (improvers – unique; deteriorators – unique). For example:

> I said that, if he wanted, we could slow things down a bit. I said it was really up to him how the sessions were inducted [sic] as he was in control. (Therapist; Simon, Affirmative Rebuttal)

> Linda's post-therapy Change Interview provides a qualitative account of this profoundly important therapeutic process, attributing this phenomenon to the counsellor offering the opportunity to allow the inner child back in. (Researcher; Linda, Affirmative Brief)

Sub-category 2.2.1.3: Being transparent. There appeared to be three main ways in which therapists demonstrated their capacity for *being transparent* (improvers – typical; deteriorators – variant) within the therapeutic process: *responding to client's questions* (improvers – unique; deteriorators – variant), *discussing process* (improvers – variant; deteriorators – variant), *noticing changes* (improvers – variant), and *expressing concern for the client* (improvers – unique). For example:

> I asked the therapist how I was doing. This was helpful as my health problems cause problems with my energy and sometimes I'm not so positive - it helped to hear what the therapist had to say. (Client; Simon, Affirmative Brief)

> The first session she was like "oh I feel like we go round in circles" and I went "oh do you mean we go round in tangents", she just said she wanted me more to focus on something instead of jumping from incident to problem to you know feelings, which has been good. I mean, I feel like, with therapist 2, I feel like there is more of a tactic going on, there's more of a kind of process going on, like I don't feel like I just came in here to talk and then walk away. (Client; Joseph, Affirmative Rebuttal)

Sub-category 2.2.2: Interfering therapist processes. The analysis revealed that some therapists experienced challenges to their capacity as person-centred therapists during the therapeutic process. I have described them as *interfering therapist processes* and noticed them particularly present among the therapists of the deteriorators (improvers – unique; deteriorators – typical). I have organised these challenges into three sub-categories: *not working in the client's frame of reference / acting in a controlling/directive manner* (deteriorators – typical), *therapist doubts* (improvers – unique; deteriorators – variant), and *inconsistent approach* (deteriorators – unique).

Sub-category 2.2.2.1: Not working in the client's frame of reference / acting in a controlling/directive manner. There were examples in the data from which it could be inferred that therapists were *not working in the client's frame of reference / acting in a controlling/directive manner* (deteriorators – typical) by directing content, attempting to accelerate process, and losing patience with their clients. For example:

> Yeah, the delving into […] my family are my family and it's like I see them often enough, I don't have to come here and talk about them for an hour because that's not what's upsetting me, that's not why I'm here, because of my family. I feel like I have a good relationship with my family. (Client; Joseph, Affirmative Rebuttal)

> On multiple occasions, Luke's therapists wrote about creating silence in the session as an "opportunity" for Luke to engage. Luke himself did not feel like this was the therapist kindly offering him a chance to engage – he viewed it as an abandonment and a lack of caring. The second therapist wrote: "He said it was not an opportunity. He said I wasn't interested." Luke described the silence as hindering on a couple of occasions. (Researcher; Luke, Affirmative Brief)

Sub-category 2.2.2.2: Therapist doubts. Therapists working with clients in both groups expressed *doubts* or behaved as if they felt doubtful during the process

(improvers – unique; deteriorators – variant). These doubts could be sorted into the following themes: *unsure how/if helping client* (improvers – unique; deteriorators – unique), *therapist at a loss to know what to do* (deteriorators – variant), and *querying their connection with the client* (improvers – unique; deteriorators – unique). For example:

> She did say at one point, "oh I'm not sure how I can help you" coz I think she felt (pause) I don't know, it was more like I was articulating in such a way that she didn't need to jump in so therefore she felt that (hesitation) a bit hindered. (Client; Joseph, Affirmative Brief)
>
> When I ask a question, the answer doesn't seem to match. Feels very 'chat' like. I wonder if I am enough – [if] what we are doing is helpful. (Therapist; Simon, Affirmative Rebuttal)

Sub-category 2.2.2.3: Inconsistent approach. There was strong evidence in the case of Luke, that therapist doubts led to an *inconsistent approach*. He worked with two therapists who each struggled in their work with him in similar but different ways. In her Affirmative Brief, the HSCED researcher (Love, 2018) argued that the evidence showed that both therapists experienced a repeated cyclical process of acceptance, frustration, unsuccessful attempts to motivate the client, withdrawal then re-engagement. This was most clearly expressed in the use by the first therapist of the threat of ending:

> The possibility of ending seemed more like an attempt to motivate Luke than a serious concern that he couldn't benefit from therapy. In the notes for session 14, the therapist wrote under 'ideas for next time': focusing on the possibility of ending to encourage him to engage. Once there was a possibility that it had 'worked', and Luke was ready to engage, therapy was offered again (though an ending was suggested again). (Researcher; Luke, Affirmative Brief)

In contrast, the second therapist's inconsistency was expressed in the range of methods by which she tried to engage the client:

> The second therapist demonstrated this by changing her approach week to week and eventually resorting to bringing in games. The games, explained in notes from sessions 41-43, appeared to initially bemuse then frustrate Luke and he referred to them as hindering or unhelpful in his last change interview. (Researcher; Luke, Affirmative Brief)

Category 2.3: Client in the Process. Once again, I found that it was possible to organize descriptions of the *client in the process* into two sub-categories: *facilitative client processes* (improvers – general; deteriorators – general) and *interfering client processes* (deteriorators – general). As you can see, this structure highlighted both similarities and differences between the two groups.

Sub-category 2.3.1: Facilitative client processes. My analysis resulted in the identification of three *facilitative client* processes. These were: *client commitment* (improvers – general; deteriorators – typical), *experiencing feelings (opening up / allowing self-awareness)* (improvers – general; deteriorators – variant) and *working through complex situations* (improvers – typical; deteriorators – variant).

Sub-category 2.3.1.1: Client commitment. There was clear evidence of *client commitment* in both groups (improvers – general; deteriorators – typical). This was indicated by *engagement in therapeutic process* (improvers – general; deteriorators – typical), *integrating therapy and life* (improvers – general; deteriorators – variant), and *working really hard in and out of the therapy room* (improvers – general).

Sub-category 2.3.1.1.1: Engagement in therapeutic process. There was strong evidence of *engagement in therapeutic process* within both groups (improvers – general; deteriorators – typical). For example:

> During the course of therapy Julia was fully engaged from the beginning which is indicated through quantitative and qualitative data. (Researcher; Julia, Affirmative Brief)

> It is notable that James engaged fully in therapy from the start and his progress is indicated qualitatively and quantitatively. (Researcher; James, Affirmative Brief)
>
> This was shown in his consistent attendance, his requests for appointment changes, over cancellations, and his willingness to financially invest in therapy. (Researcher; Luke, Adjudication)

Sub-category 2.3.1.1.2: Integrating therapy and life. A capacity to *integrate therapy and life* appeared to be a common feature for improvers but less so for deteriorators (improvers – general; deteriorators – variant). This was exhibited through evidence of them *using therapy to connect with / work on life problem* (improvers – general; deteriorators – variant), which for some included *using therapy to gain distance / perspective on life problem* (improvers – variant) or *reflecting on progress in life problem* (improvers – unique). For example:

> Talking about getting off my anti-depressants had me feeling really anxious about it in the session, but then we turned it around and I was reminded that I do know how to ground myself in an anxious situation. (Client; Julia, Affirmative Brief)
>
> I realized that I have made some progress in integrating different aspects of my life/self into a whole. This helps me to interact with other people with less trepidation. (Client; James, Affirmative Brief)

Sub-category 2.3.1.1.3: Working really hard in and out of therapy room. In addition, there was evidence that improvers were *working really hard in and out of the therapy room* (improvers – general) including *self-help / self-initiated efforts*. For example:

> In her six month follow-up change interview, Julia clearly reports that it's work done inside and outside of therapy that helped her: "I think being better off now came from really hard work in here but also really hard work outside of the therapy room". (Researcher; Julia, Affirmative Rebuttal)

> Most of Linda's therapist's identified Extra-Therapy Events relate to material that Linda found helpful working through in therapy; acting as situational or experientially specific examples of broader issues that Linda was able to change. For example, in Session 9, Linda's distressing interaction with her husband may have initiated or progressed Linda's ability to explore and work through this ongoing issue in therapy. (Researcher; Linda, Affirmative Rebuttal)
>
> James was preparing to move when he began therapy and saw this move as a metaphor for a new beginning. In the initial therapy sessions, the client seemed to equate the idea of moving house and having a fresh start with his hopes of making personal change at the start of therapy. (Researcher; James, Sceptic Brief)

Sub-category 2.3.1.2: Experiencing feelings (opening up / allowing self-awareness). Clients in both groups described a process of *experiencing their feelings, (opening up / allowing self-awareness)* during therapy but this was particularly evident for improvers (improvers – general; deteriorators – variant). In particular, there were more detailed descriptions of improvers *realizing feelings* (improvers – typical), *working with / resolving / releasing stuck feelings* (improvers – variant; deteriorators – unique), *greater general awareness of feelings* (improvers – variant), and *valuing feelings* (improvers – unique).

Sub-category 2.3.1.2.1: Realizing feelings. Typically, the therapeutic process enabled improvers to begin *realizing their feelings* (improvers – general). For example:

> Talking about it makes me realize some feelings I couldn't label before. My counsellor has a really good way of putting words to it. (Client; Julia, Affirmative Brief)
>
> Starting to deal with the feelings of shame and guilt is the deepest most important breakthrough in the therapy so far. The realization and new understanding of these issues allows me to find ways of

> finally tackling/dealing with the problem. (Client; James, Affirmative Brief)

Sub-category 2.3.1.2.2: Working with / resolving / releasing stuck feelings. Some clients (improvers – variant) described a process of *working with / resolving / releasing stuck feelings* during their therapy. In particular, this was experienced as *letting go* (improvers – variant) and *letting in (inner child)* (improvers – variant).

> I felt that it was safe to share/express my feelings in a more emotional way and not just in an academic way. This is quite liberating and a relief to be able to unburden some of the stuff that has been weighing me down. (Client; James, Affirmative Brief)

> A specific therapy process that Linda identified as being extremely helpful was the exploration or 'letting in' of her inner child, which she described as making her feel complete. This process was identified as a profound moment of relational depth between Linda and her therapist. (Researcher; Linda, Affirmative Brief)

Sub-category 2.3.1.2.3: Greater general awareness of feelings. Two improvers noted that they experienced a *greater general awareness of feelings.* For example:

> It's like, I'm more aware of where my feelings are coming from. (Client; Linda, Affirmative Brief)

Sub-category 2.3.1.2.4: Valuing feelings. One improver developed a real *valuing of his feelings* during the therapy process:

> I suppose feeling a sense of anger informs me of what more power/strength I might have to use rather in a more positive way. It was kind of empowering. (Client; James, Affirmative Brief)

Sub-category 2.3.1.3: Working through complex situations. There was evidence of clients using the therapeutic process to *work through complex*

situations (improvers – typical; deteriorators – variant). They did this through *exploring* (improvers – typical; deteriorators – variant) and by *making connections / understanding why things happened* (improvers – typical).

Sub-category 2.3.1.3.1: Exploring. The process of *exploring* (improvers – typical; deteriorators – variant) enabled clients in *sorting thoughts more clearly* (improvers – typical; deteriorators – unique), *articulating thoughts* (improvers – unique; deteriorators – variant), *identifying main issues* (improvers – unique; deteriorators – unique) and *looking at issues in depth* (improvers – unique; deteriorators – unique). For example:

> An aspect which stands out in these forms regarding how Caitlin experiences therapy with both therapists is being able to explore the content she brings. (Researcher; Caitlin, Sceptic Brief)

> My method works for me and that was what was good with therapist 1 and therapist 2, I feel like sometimes I might be talking complete nonsense, but I get to vocalise and then I go away and think oh actually I do do that. (Client; Joseph, Sceptic Brief)

> Managed to identify more or less the main different issues that I am dealing with in life. (Client; James, Affirmative Brief)

Sub-category 2.3.1.3.2: Making connections / understanding why things happened. Some clients described a process of *making connections / understanding why things happened* (improvers – typical. This included *recognizing what issues are in the past* (improvers – unique) or *seeing current circumstances in larger context* (improvers – unique):

> I have been looking for the (why) answer to my marriage break up. I feel free. (Client; Linda, Affirmative Brief)

> There has just been progressively more understanding of how all the different issues in my life fit together. (Client; James, Affirmative Brief)

> This helped me to decide to look at my life and decide that some things are in the past. (Client; Simon, Affirmative Brief)

Sub-category 2.3.2: Interfering client processes. There was evidence that all four deteriorators, but none of the improvers, experienced some form of *interfering client process*. I have organized this data into four sub-categories: *discomfort with process* (deteriorators – general), *deference toward therapist* (deteriorators – typical), *struggle to integrate therapy and life* (deteriorators – typical), and *lack of engagement* (deteriorators – variant).

Sub-category 2.3.2.1: Discomfort with process. All deteriorators experienced some form of *discomfort in the process* (deteriorators – general). This discomfort arose in several ways: *difficulty opening up* (deteriorators – general), *struggling to find direction* (deteriorators – typical), *feeling that they had to change / do therapy differently / felt pressure from therapist* (deteriorators – variant), and *feeling uncared for by therapist* (deteriorators – unique).

Sub-category 2.3.2.1.1: Difficulty opening up. It is clear from the data that *opening up* was experienced as *difficult and painful* (deteriorators – typical):

> The last session I had with therapist 2 was quite challenging so when I came out feeling very depressed after it, whereas that's never really happened before, I've never come out feeling upset so, but that last session I felt quite depressed by it, not because of what she said but what we talked about. (Client; Joseph, Affirmative Brief)

> Going through things from childhood has been difficult and painful. (Client; Caitlin, Affirmative Brief)

For some, this may have been because they felt a *pressure to engage* (deteriorators – typical). For example:

> Sofia reports struggling to come to terms with the changes she made, therefore this brief suggests that the therapist was directive and

> perhaps pushed Sofia too hard to discuss or explore issues she was not ready to. (Researcher; Sofia, Affirmative Brief)
>
> He felt great pressure to engage, as evidenced by his anxiety at the start of session 22 after an ending had been suggested again the previous session; if he did not behave in a certain way, this relationship would end. (Researcher; Luke, Affirmative Brief)

Other clients struggled to open up because they were *feeling exposed* (deteriorators – variant) or *concerned about upsetting the therapist* (deteriorators – unique):

> Opening up but with no resolution – left feeling intense vulnerability. (Adjudication panel; Caitlin, Adjudication)
>
> And then, when I upset her, I was like, "oh god, if this had carried on, would there have been more of that, would I have been having to hold back so I didn't upset my counsellor?" (Client; Joseph, Affirmative Brief)

Sub-category 2.3.2.1.2: Struggling to find direction. Another aspect of the discomfort experienced by several clients was *struggling to find direction* (deteriorators – typical). For some, this took the form of *struggling to know what to say* (deteriorators – unique), noticing that therapy *can feel like a chore to find useful direction for work* when they had a good day (deteriorators – unique) and *coasting* (deteriorators – unique). For example:

> He described finding the lack of direction from the therapists difficult and that he struggled with setting his own direction in the therapy, adding that he would prefer to be given goal or targets and for advice to be given. (Researcher; Joseph, Sceptic Brief)
>
> Luke frequently brought no content into his sessions. This was due to the fact that he struggled knowing what to say. (Researcher; Luke, Sceptic Rebuttal)

> And it feels like, I'm kind of like, eyeing the clock to see how much longer I can talk about where if I was being honest, I'd go, can we just wrap this up early coz I don't, but then I know there's a whole process to go through. (Client; Joseph, Affirmative Rebuttal)

Sub-category 2.3.2.1.3: Feeling that they had to change / do therapy differently / felt pressure from therapist. There was evidence that two clients experienced discomfort due to *feeling that they had to change / do therapy differently / felt pressure from therapist*, for example:

> When she said "I don't know how we can help you", I felt like, I don't know, does that mean I'm completely unhelp-able, I'm beyond saving or I'm being very unhelpful or obtuse - that kind of thing […] I didn't take it personally, I wasn't upset by it, but it was more like, "oh maybe what I'm saying, I have to change how I say it". (Client; Joseph, Affirmative Rebuttal)

Sub-category 2.3.2.1.4: Feeling uncared for by therapist. For one client, his discomfort in the process was worsened by their perception that their therapist did not care about how difficult it was for him:

> Luke described feeling uncared for by his second therapist; for him, it was a massive challenge to participate in therapy. In his fourth change interview, Luke said that his therapist seemed indifferent to his struggle in finding something to say. (Researcher; Luke, Affirmative Brief)

Sub-category 2.3.2.2: Deference toward therapist. There was evidence that suggested that some clients (deteriorators – typical) developed a *deference toward their therapist*, which may have prevented them from getting what they needed from the therapeutic process. For example, it is possible to infer from the data that clients may have deferred to their therapists' power in the relationship, whether as experts in the process or in respect of their availability for sessions:

> The TRS scores suggest that the therapist took the lead in sessions and that Sofia felt less able to disagree with or correct her, which suggests that Sofia did not have enough support from the therapist to feel comfortable with these changes and perhaps that they were not conclusions she came to of her own accord or that she was fully accepting of. (Researcher; Sofia, Affirmative Rebuttal)

> While it cannot be denied that there were significant gaps between Caitlin's appointments, particularly with her second therapist, Caitlin notes in her Mid-4 Change Interview that these inconsistencies could not be helped. (Researcher; Caitlin, Sceptic Rebuttal)

Sub-category 2.3.2.3: Struggle to integrate therapy and life. In contrast to the improvers' apparent ability in this area, there was clear evidence of deteriorators *struggling to integrate therapy and life* (deteriorators – typical). It was clear that these clients experienced both situational barriers, such as Joseph's difficulty in attending regularly due to work commitments then experiencing a significant life event when he was waiting to transfer to a new therapist, as well as personal barriers in putting new understanding gained in therapy into action in life. For example:

> The most frequently referred to hindering personal characteristic expressed by Caitlin was that she believes herself to be stubborn and unwilling to accept or consider another individual's point of view, this was stated in her 1st, 2nd and 3rd Change Interviews. One might consider this to be a barrier to growth in therapy, particularly as Caitlin expressed at her Mid-1 Change Interview that she struggles to apply what she learns in therapy to "outside the counselling room". (Researcher; Caitlin, Sceptic Brief)

Sub-category 2.3.2.4: Lack of engagement. Finally, there was evidence of some clients' *lack of engagement* in the process (deteriorators – variant). For example:

> In his HAT forms, Luke expressed frustration with sitting in silence and having pointless conversations. While the affirmative case attributes the blame for this to the therapist, Luke himself actually attributes a proportion of the blame to himself. In his change interviews, Luke admits he "[hasn't] used the chance to interact and relate". (Researcher; Luke, Sceptic Brief)

Category 2.4: Challenges in the Process. It became clear during the analysis that some clients, deteriorators in particular, experienced *challenges in the therapeutic process* (improvers – variant; deteriorators – general). One of these challenges was *change of therapist* (improvers – variant; deteriorators – general), which had been identified when clients were selected for this study; another challenge that emerged was *delay in process* (improvers – unique; deteriorators – typical).

Sub-category 2.4.1: Change of therapist. As noted in the Method section, all four deteriorators and two of the four improvers experienced at least one *change of therapist* during their period of counselling at the research clinic. In the analysis, I found data relating to the experience of changing therapist that I organized into two sub-categories: *non-interfering effect of change of therapist* (improvers – general, i.e. all improvers who experienced a change of therapist; deteriorators – typical) and *interfering effect of change of therapist* (deteriorators – general).

Sub-category 2.4.1.1: Non-interfering effect of change of therapist. I sorted evidence of the *non-interfering effect of change of therapist* into three sub-categories: *positive perception* (improvers – general; deteriorators – general), *beneficial impact* (improvers – variant, i.e. one of two improvers who experienced a change of therapist; deteriorators – variant), and *no impact* (improvers – general; deteriorators – unique).

Sub-category 2.4.1.1.1: Positive perception. Some of the clients (improvers – general; deteriorators – general) expressed a *positive perception* about the process of changing therapist:

> It's good to be back! I can't remember when I was last here but feel positive to meet new therapist. (Client; Simon, Affirmative Brief)

For some, their positive perception arose from the opportunity it presented: they saw it as *like a new start* (deteriorators – variant) or because it was a chance to gauge if they were *handling it* (the change) *well* (improvers – general):

> Yeah, their tactics were completely different, I mean, this is why it feels like a new start, I mean working with a new person (Client; Joseph, Affirmative Rebuttal)

> I met the new therapist for the first time. I had been a bit anxious before but am now somewhat relieved. (Client; James, Affirmative Brief)

Sub-category 2.4.1.1.2: Beneficial impact. It appeared that some clients who went through a change of therapist found the process *beneficial* (improvers – variant; deteriorators – variant) because the second therapist appeared to be a better fit for the client. For example:

> And I think therapist 2 is quite good because she's a bit more of a… she'll challenge what I say or want me to pick up on something and delve into it, whereas therapist 1 was more happy to let me talk and just respond occasionally and I'm not really sure that is what I was after. I know that the point of these sessions is that I do the talking and it's… I kind of lead the conversations but therapist 2 is more happy to jump in when therapist 1 was a bit held back a bit. (Client; Joseph, Sceptic Brief)

Sub-category 2.4.1.1.3: No impact. There was also evidence that the change of therapist had *no impact* (improvers – general; deteriorators – unique) for two apparent reasons: *both/all therapists were helpful* (improvers – variant; deteriorators – unique) and that the client felt *attached to the process not the therapist* (improvers – variant):

> James' responses in the Change Record suggest that he found therapy helpful while working with both of his therapists. (Researcher; James, Affirmative Brief)

> You kinda get to a point where you're spending an hour a week in a room with someone doing absolutely amazing stuff so all you can say realistically for me is that there may be some positive attachment not to the person but with the process, and em, it's a really good experience. (Client; Simon, Affirmative Brief)

Sub-category 2.4.1.2: Interfering effect of change of therapist. Despite some having positive perceptions and experiencing potential benefits in the change - the evidence suggested that there were also *interfering effects of change of therapist* in particular for deteriorators (deteriorators – general). Indeed, the adjudication panels for three of the four deteriorators (Joseph, Sofia, Caitlin) identified the change of therapist as hindering and a potential contributing factor in their deterioration. My analysis identified several ways in which the change of therapist appeared to be interfering: *unfortunate timing for client* (deteriorators – typical), caused *disruption in building therapeutic relationship* (deteriorators – variant), *reinforcing expectation of being rejected in relationship* (deteriorators – unique), and *difficulties in relationship with second therapist* (deteriorators – unique).

Sub-category 2.4.1.2.1: Unfortunate timing for client. It was apparent from the data that the change of therapist was *unfortunate timing for the client* (deteriorators – typical), either because the client was *in mid-process* (deteriorators – variant) or because they experienced it as a *loss of support when feeling vulnerable* (deteriorators – unique). For example:

> This evidence presented here suggests that the change of therapist may have been very difficult for Sofia. This is of particular importance if this was a time when Sofia was under a lot of pressure to make a very important decision about her future. (Researcher; Sofia, Affirmative Brief)

> Between sessions 26-31 when she was undergoing frequent changes to her medication, Caitlin also had a change of therapist at this time (session 26). She highlights both of these changes as hindering in her Mid-3 Change Interview. These hindering factors with the additional deterioration in her Mid-3 PQ items: 'I feel overwhelmed by my depression' and 'Sometimes I do not feel like I get much support', suggest that the change of therapist was untimely for Caitlin and did not aid her in feeling fully supported. (Researcher; Caitlin, Affirmative Rebuttal)

Sub-category 2.4.1.2.2: Disruption in building therapeutic relationship. The data showed that the *disruption in building the therapeutic relationship* caused by the change of therapist was particularly difficult for some clients (deteriorators – variant):

> It is clear from the change interview that Joseph struggled with the changes in therapist and that he found it difficult to build relationships. Given the amount of sessions that he attended (16 in total) it is not unreasonable to assume that the majority of these were taken up with either establishing the relationship or ending it. (Researcher; Joseph, Affirmative Brief)

> There is substantial evidence to prove a positive relationship with the first therapist, the deterioration occurs after Sofia changed therapist. It can be difficult build a connection after a transfer, especially when there have been a substantial number of sessions with the first and that a strong therapeutic relationship has been built. (Researcher; Sofia, Affirmative Rebuttal)

Sub-category 2.4.1.2.3: Reinforcing expectation of being rejected in relationship. The data suggested that the change in therapist had a twofold interfering effect on one deteriorator. First, because his perception of being

rejected by his first therapist had the impact of *reinforcing his expectation of being rejected in relationship:*

> In session 18, the [therapist] notes read: 'I said that I think the next session will be our last. He said that he expected it to go this way however when asked to elaborate he chose not to explain what he meant.'[…] Luke's first PQ item, "my relationships fall apart soon after they begin", suggests that Luke expected to be rejected by the therapist, since that is how all his other relationships ended. (Researcher; Luke, Affirmative Brief)

Sub-category 2.4.1.2.4: Difficulties in relationship with second therapist. The change of therapist continued to have impact on this client because he experienced *difficulties in relationship with his second therapist:*

> In fact, the TRS data suggests that overall Luke's relationship with the second therapist was worse. Losing his temper with her does not mean that he would not have done the same with the first therapist, had he done something to make Luke angry. The change in expressed emotion coincides with the change in therapist, therefore it cannot be assumed that it had anything to do with a change within Luke. It is very possible that the change in Luke was his reaction to a therapist who he perceived to be finicky and unempathetic. (Researcher; Luke, Affirmative Rebuttal)

Sub-category 2.4.2: Delay / inconsistency in process. There was evidence that some clients experienced a *delay / inconsistency in the process* (improvers – unique; deteriorators – typical) for two main reasons: *inconsistency of sessions* (improvers – unique; deteriorators – variant), and because it *took time to develop the relationship* (deteriorators – variant).

Sub-category 2.4.2.1: Inconsistency of sessions. Three clients were affected by *inconsistency of sessions* (improvers – unique; deteriorators – variant). It can be

inferred from the data that this inconsistency created a *sense of abandonment* for some of these clients (deteriorators – variant):

> And then when that thing happened with my personal situation I really needed it but it wasn't available […] I sort of found it annoying. (Client; Joseph, Affirmative Brief)

> I would prefer if my counselling is more stable though, as my anxiety goes through the roof when I'm unsure about things. (Client; Caitlin, Affirmative Brief)

Sub-category 2.4.2.2: Took time to develop relationship. There was evidence that some relationships were slow to develop: a *'rocky start'* (deteriorators – unique) or *the need to develop common understanding* (deteriorators – unique). For example:

> She had a rocky start in therapy and struggled to make a connection to her therapist in the first session. (Researcher; Sofia, Affirmative Narrative)

> It took me a while to get anywhere with her because she didn't seem to be, not that she was baffled by what I was saying but her personal life experience was completely different to mine, she married with children or divorced with children and I don't have that kind of lifestyle, it felt that we were trying to get a common understanding, trying to get to know each other. (Client; Joseph, Affirmative Rebuttal)

In the Process: Summary. The general and typical findings in this domain for each group are presented in Table 6.4 and summarized below. Any finding which is different to the other group by at least two frequency counts is highlighted in bold.

Both improvers and deteriorators experienced the *therapeutic climate* as *safe*, containing *warmth and connection*, and *supportive*, enabling *facilitative therapist processes* and *facilitative client processes* to occur in the therapy.

Table 6.4. Domain 2: In the Process

Category	Improvers	Deteriorators
Therapeutic climate	*Safe* *Warmth / connection* *Supportive*	*Safe* *Warmth / connection* *Supportive*
Therapist in the process	Facilitative therapist processes: Accepting/ understanding/ validating client -> *working within client's frame of reference*; *Offering opportunities;* *Being transparent.*	Facilitative therapist processes: Accepting/ understanding/ validating client -> ***validating/ affirming client's experience*** ***Interfering therapist processes:*** ***Not working in client's frame of reference/ acting in a controlling/ directive manner***
Client in the process	Facilitative client processes: Client commitment -> engagement in therapeutic process; **integrating therapy & life -> using therapy to connect with/ work on life problem; working really hard in & out of therapy room;** **Experiencing feelings (opening up/ allowing self-awareness) -> *realising feelings;*** *Working through complex situations -> exploring ->* ***sorting thoughts more clearly; making connections / understanding why things happened***	Facilitative client processes: *Client commitment -> Engagement in therapeutic process* **Interfering client processes:** **Discomfort with process -> difficulty opening up -> d*ifficult & painful -> pressure to engage?; Struggling to find direction;*** ***Deference toward therapist;*** ***Struggle to integrate therapy & life***
Challenges in the process	Change of therapist -> Non-interfering effect -> **positive perception -> handling it well; no impact**	Change of therapist -> Non-interfering effect; **interfering effect -> *unfortunate timing for client;*** ***Delay/ inconsistency in process***

Notes. **Bold** text = difference of at least two frequency counts; non-italicized text = general frequency; italicized text = typical frequency.

Their therapists were *accepting / understanding / validating of the client* but the data suggests that there may have been different emphases. It was typical for

improvers' therapists to demonstrate acceptance of their clients by *working within the client's frame of reference*. The deteriorators' therapists did so by *validating and affirming the client*. This was less common for improvers' therapists, highlighting an apparent difference between the two groups. In addition, there was evidence that it was typical for the improvers' therapists to be facilitative by *offering opportunities* and *being transparent*; this was not a typical finding for the deteriorators' therapists but there was sufficient data indicating its presence for this not to be a finding of difference between the two groups. In contrast, I found difference in relation to *interfering therapist processes.* In particular, there was evidence inferring that it was typical, at times, to find that the deteriorators' therapists were *not working in the client's frame of reference or acting in a controlling/directive manner*. This suggests that these therapists struggled more to maintain a person-centred attitude when working with these clients than the therapists who worked with the clients who improved by the end of therapy.

This difference in the therapists' experience is reflected in the descriptions of the clients in process that have emerged from this analysis. Again, both groups of clients were able to access *facilitative client processes.* There was clear evidence of *client commitment* present in both groups, in particular *engagement in the therapeutic process.* However, there were two clear differences between the improvers and deteriorators. One was their apparent ability to *integrate therapy and life*. Improvers were able to do this: the data showed that they *used their therapy to connect with and work on their problems.* In addition, it was clear that they *worked hard both in and out of the therapy room*. Improvers described a process in therapy of *experiencing their feelings that involved opening up and allowing self-awareness* and of *working through complex situations*, in particular by *exploring,* including *sorting thoughts more clearly*, and *making connections / understanding why things happened*. There was less evidence of deteriorators describing their process in this detail.

In contrast, deteriorators experienced *interfering client processes* that improvers did not seem to have had to contend with. These included: *discomfort*

with the process, in particular finding it *difficult and painful to open up* - perhaps feeling a *pressure to engage* - and *struggling to find direction*; *deference toward the therapist;* and a *struggle to integrate therapy and life*.

The analysis identified that deteriorators seemed to experience more interfering effects arising from *challenges in the therapeutic process* than the improvers. Although two of the four improvers experienced a *change of therapist* at least once during their therapy, the data suggested that they perceived this change positively and that it had no particular impact on them. In contrast, the *change of therapist* appeared to have had both *non-interfering* and *interfering effects* on the deteriorators, in particular because it happened with *unfortunate timing for the client*. Lastly, there was evidence that deteriorators were more affected by *delay / inconsistency in the process* than improvers.

Domain 3: In the End

The final domain presents an analysis of the data contained in the HSCED material that focused on the clients at the end of their therapeutic process, organized into four categories: *ending of therapy*, *changes* and *potential impact of research*.

Category 3.1: Ending of therapy. As clients were designated improvers and deteriorators according to their scores on the Strathclyde Inventory at the end of therapy, it seemed important to look at any evidence contained in the data about the context in which this *ending of therapy* took place. In this category, I organized into two sub-categories: *facilitative aspects of ending* (improvers – general) and *interfering aspects of ending* (deteriorators – general). Doing so highlighted a stark contrast in the experience of the end of therapy between the two groups.

Sub-category 3.1.1: Facilitative aspects of ending. The data suggested that there were three *facilitative aspects of ending* experienced by clients in the study (improvers – general): *feeling ready to end* (improvers – general), perceiving that the *sessions helped a lot* (improvers – general), and reaching a *mutual decision to*

end (improvers – typical). There was no data to suggest that any of these facilitative aspects of ending were experienced by any of the deteriorators.

Sub-category 3.1.1.1: Feeling ready to end. There was evidence that all improvers were *feeling ready to end* their therapy. Some stated explicitly that they felt *able to continue the progress made (improvers – variant)*. For example:

> Letting go of sessions is scary but being able to, feels like great progress. (Client; Julia, Affirmative Brief)
>
> Today was the end of our sessions – I feel it is time – I feel whole and ready to embrace life. (Client; Linda, Affirmative Brief)
>
> I suppose that the process of ending was handled in a good way which was helpful for me to deal with. As previously I have felt quite ambivalent about endings. But now I feel like I am successfully moving forward onto the next stage after having accomplished a lot. Progress has been made. Very beneficial! (Client; James, Affirmative Brief)
>
> It felt like the natural end of therapy. I feel more independent - I talked about more of happiness + health. I feel I respect the process involved + I carry it with me + already today something shifted for me before I got here. It feels like it's part of my nature. (Client; Simon, Affirmative Brief)

Sub-category 3.1.1.2: Sessions helped a lot. All improvers believed that their therapy *sessions helped a lot*. This belief was supported by their perceptions that they had *made great progress* (improvers – typical), or that they *felt really good after therapy* (improvers – variant). For example:

> These sessions have helped a lot, often things have greatly shifted. (Client; Julia, Affirmative Brief)
>
> I am at the point in my life where I have made massive progress + recovered with some issues still to be worked on + that's so good

> that I can now steer through the positive + that I can value the time, place, and people, and resources I have in my life. Not everyone has access to therapy. (Client; Simon, Affirmative Brief)

Sub-category 3.1.1.3: Mutual decision to end. For three improvers, there was a *mutual decision to end* reached in discussion with their therapist. (The fourth improver, James, ended because he had reached the maximum number of sessions available to him.)

> We agreed to end after our next session, and both acknowledged that something huge had happened that couldn't be explained. (Therapist; Linda, Affirmative Brief)

> Talking about where we are, working towards 45 sessions. (Therapist; Simon, Affirmative Rebuttal)

Sub-category 3.1.2: Interfering aspects of ending. In contrast, all deteriorators experienced *interfering aspects of ending.* These can be summarized as an *incomplete therapeutic process* (deteriorators – general) characterized by *client decision not to continue following therapist's decision to leave / take extended break* (deteriorators – typical), the client *did not engage in ending process* (deteriorators – variant) or *left feeling distressed and vulnerable* (deteriorators – variant).

Sub-category 3.1.2.1: Incomplete therapeutic process. All deteriorators experienced an *incomplete therapeutic process* (deteriorators – general). For example:

> Therefore, at the time that therapy ended, Sofia was not in a good place and seemed to get very little from her five final therapy sessions. She had not fully worked through some very important personal issues. In fact, therapy had brought up a lot for her, particularly in relation to her experience of life in the UK. She had not identified what it was that she felt she was missing in the UK

> compared to life back home. It seems like Sofia ended her time in therapy too soon. (Researcher; Sofia, Affirmative Narrative)

Sub-category 3.1.2.1.1: Client decision not to continue following therapist's decision to leave / take extended break. In contrast to the improvers' experience of reaching a mutual decision to end, the data suggested that for most deteriorators the therapy ended when they made the *decision not to continue following their therapist's decision to leave / take an extended break* (deteriorators – typical). For example:

> A further element reported by the therapist to have potentially caused distress is a further change of therapist which was disclosed to Caitlin at session 44; her therapist wrote that she 'Seemed slightly disappointed about having to change counsellor in the new year.' (Researcher; Caitlin, Affirmative Brief)
>
> At the point at which Joseph left therapy he would have been allocated a fourth therapist should he have decided to remain. (Researcher; Joseph, Affirmative Brief)
>
> I remember that I said to you that I will be back in mid-August and we said that you will get in touch to restart sessions if you want them. I said that I will keep a place for you till the start of September. […] I am aware that if you do not come back, that I may not see you again (Therapist; Luke, Sceptic Rebuttal)

Sub-category 3.1.2.1.2: Did not engage in ending process. Some clients chose *not to engage in an ending process* with their therapist or the research clinic. For example:

> At the time when Joseph ceased therapy he did not engage in any ending or completion of therapy and did not attend for his end of therapy interviews, suggesting that he felt a further change of therapist was not helpful to him and that he felt that therapy was not helpful to him. (Researcher; Joseph, Affirmative Brief)

> They ended therapy [by email when therapist began extended break] on a very warm note, laughing together and showing regret that they didn't end properly. (Researcher; Luke, Sceptic Rebuttal)

Sub-category 3.1.2.1.3: Left feeling distressed and vulnerable. There was data that inferred some clients *left therapy feeling distressed and vulnerable* (deteriorators – variant). For example:

> Sofia's discomfort at living in the UK seems to become more prevalent towards the end of therapy. The therapist notes from her final session indicate that the client discusses her dilemma over moving home or staying in the UK [and] that that Sofia is upset as this decision has been impacting negatively on her relationship with her partner. (Researcher; Sofia, Affirmative Brief)

> Client got upset about ending counselling at the research clinic. (Therapist; Caitlin, Affirmative Brief)

This finding will be presented in more detail in the *Changes* section.

Category 3.2: Changes. In my analysis I sorted the *changes* experienced by clients by the end of therapy into two sub-categories: *facilitative changes* (improvers – general; deteriorators – general) and *interfering changes* (improvers – unique; deteriorators – general).

Sub-category 3.2.1: Facilitative Changes. All clients experienced changes that could be inferred from the data to be *facilitative changes*. I organized these changes into four sub-categories: *improvement in key difficulties that wanted to address in therapy* (improvers – general; deteriorators – general), *relationship with self* (improvers – general; deteriorators – general), *self in the world* (improvers – general; deteriorators – general), and *positive change in personal circumstances* (improvers – typical; deteriorators – unique).

Sub-category 3.2.1.1: Improvement in key difficulties that wanted to address in therapy. There was evidence that both improvers and deteriorators experienced

improvement in key difficulties (improvers – general; deteriorators – general). For some, this improvement occurred *during early stages in therapy* (improvers – typical).

> Positive change was reported on all of the [PQ] items and ranged from 1 to 2 points. The drastic change is quite notable. (Researcher; Julia, Affirmative Brief)
>
> Two out of the ten individual items were removed from the Personal Questionnaire at Simon's request within the first ten sessions. At intake both items were rated 6 by Simon which equates to a very considerable level of impact up until the point he asked for them to be removed. (Researcher; Simon, Affirmative Brief)
>
> An indication that Caitlin did experience positive change through attending therapy can be observed firstly in her improved PQ score for her item "I have trouble controlling my anger" which was added at her Mid-2 Change Interview. This item when added was rated as having bothered Caitlin "very considerably", and at the end of therapy this had changed to "very little". (Researcher; Caitlin, Sceptic Brief)

Sub-category 3.2.1.2: Relationship with self. My analysis found evidence to suggest that all clients experienced some degree of change in their *relationship with self.* I was able to organize the data into three further sub-categories: *increased self-awareness / self-understanding* (improvers – general; deteriorators – general), *increased self-acceptance* (improvers – general; deteriorators – variant), and *feeling complete / more integrated* (improvers – general; deteriorators – variant).

Sub-category 3.2.1.2.1: Increased self-awareness / self-understanding. Both improvers and deteriorators demonstrated *increased self-awareness / self-understanding* by the end of therapy. For example:

> It's like, I'm more aware of where my feelings are coming from. Or where things... I'm tending to look at things and think well that's

> because of that. If this had happened, that had happened. (Client; Linda, Affirmative Brief)
>
> In Luke's second change interview, when asked to sum up what made therapy helpful, Luke stated that he had become more aware. (Researcher; Luke, Sceptic Rebuttal)

Most clients became *more aware of their needs and their impact* (improvers – general; deteriorators – typical):

> Realizing that where being a little hard on myself used to motivate me, it only makes me feel worse now that the reason I'm being hard on myself is more personal. (Client; Julia, Affirmative Brief)
>
> I realized that guilt and shame are holding me back and that this guilt and shame is not rational even though they were once useful as a defense mechanism. (Client; James, Affirmative Brief)
>
> Mid-4 Change: Client acknowledges that she is deserving of support and notes this as "surprising". (Researcher; Caitlin, Affirmative Brief)

Sub-category 3.2.1.2.2: Increased self-acceptance. There was evidence that some clients, in particular improvers, developed an increased capacity for *self-acceptance* by the end of therapy (improvers – general; deteriorators – variant). This *self-acceptance* was characterized by *more self-compassion* (improvers – general; deteriorators – variant) and *increased appreciation of self* (improvers - typical):

> Realizing that I have become better at allowing myself to have off moments. Better at self-compassion. (Client; Julia, Affirmative Brief)
>
> James gained greater self-acceptance. "It's okay to be me and where and how I am […] I feel freer to be me, freer to be happy. I don't… I think somehow I used to feel that I oughtn't to be happy. I had a lot of shame and guilt, um, which was really unwarranted." (Researcher; James, Affirmative Brief)

> This felt good because I was able to flow a bit more than usual + used time wisely as I have with other aspects of my life. (Client; Simon, Affirmative Brief)

Sub-category 3.2.1.2.3: Feeling complete / more integrated / back on track. The final way in which clients expressed changes in their relationship with themselves was in describing themselves as *feeling complete / more integrated / back on track* (improvers – general; deteriorators – variant). In particular I found descriptions of clients experiencing *increased self-control / stillness / calm* (improvers – typical; deteriorators – unique), *more access to emotions* (improvers – unique; deteriorators – unique), and *more in touch with embodied self* (improvers – unique).

> I feel like I am back to me. (Client; Julia, Affirmative Brief)

> One of Linda's most distressing PQ Items "I feel out of control" (as identified in her top 3 most distressing items), also relates distinctly to changes throughout therapy, with Linda developing an awareness of her own feelings and processes, which allows her to regain control of her reactions, and become stiller. (Researcher; Linda, Affirmative Brief)

> My body feels better having me around, I think, it's less, my body feels less anxious, less uncertain, abandoned. I'm spending more time though in my body which is a good thing cuz I used to think most of my time in some part of my head which was connected to outer space. My body feels less anxious, less uncertain, abandoned. I feel happier now that I'm more integrated in my body. (Client; James, Affirmative Brief)

Sub-category 3.2.1.3: Self in the world. There was evidence of changes in relation to their *self in the world* for clients in both groups (improvers – general; deteriorators – general). I was able to organise the data supporting this sub-category into four further sub-categories: *more empowered / motivated* (improvers

– general; deteriorators – variant), *new understanding of self in relationship* (improvers -typical; deteriorators – variant); *increased openness* (improvers – general; deteriorators – unique), and *more positive / balanced /realistic perspective on life* (improvers – general).

Sub-category 3.2.1.3.1: More empowered / motivated. There was data from six clients (improvers – general; deteriorators – variant) that suggested they felt *more empowered / motivated* at the end of therapy.

> Piecing things together has an illuminating and empowering effect. (Client; James. Affirmative Brief)
>
> Now I feel I want to move forward and now that my head and heart is clearer- make a detailed plan – and take action. (Client; Simon, Affirmative Brief)

Most of these clients (improvers – typical; deteriorators – unique) described having *increased self-confidence / self-trust / self-belief*. For example:

> I am more confident in public speaking and social interaction. (Client; Julia, Affirmative Brief)
>
> What made this important/significant is that I believe in myself more. (Client; Simon, Affirmative Brief)
>
> She gained confidence through the therapeutic alliance (Adjudication panel; Sofia, Adjudication)

Sub-category 3.2.1.3.2: New understanding of self in relationship. There was evidence of clients leaving therapy with *new understanding of self in relationship* (improvers – typical; deteriorators – variant). I organized this data into four sub-categories: *less fearful of rejection and abandonment* (improvers – typical; deteriorators – unique), *more able to put own needs first in relationship* (improvers – variant; deteriorators – unique), *increased awareness of impact on/of self in relationship* (improvers – variant), and *figuring out who/how to trust* (improvers – unique).

Sub-category 3.2.1.3.2.1: Less fearful of rejection and abandonment. There was evidence that some clients experienced change in their relationships through becoming *less fearful of rejection and abandonment* (improvers – typical; deteriorators – unique), for example, by becoming *able to deal with conflict in relationship* (improvers – unique; deteriorators - unique):

> According to Linda's PQ duration, her longest-standing problem (6-10 years) was Item 7: "I feel that I fear rejection and abandonment". Linda's Change Interview Record showed positive changes throughout therapy that could relate to this item. (Researcher; Linda, Affirmative Brief)
>
> Feeling better equipped to deal with conflicts. (Client; Caitlin, Sceptic Brief)

Sub-category 3.2.1.3.2.2: More able to put own needs first in relationship. Another indicator of change in clients' relationships with others was evidence that they become *more able to put own needs first* (improvers – variant; deteriorators – unique). This was illustrated in several ways, such as being *able to ask for needs to be met* (improvers – unique), *setting limits* (improvers – unique) and being *able to make choices for self* (deteriorators – unique). For example:

> James feels freer to put his needs first over the needs of others, [saying] "…a responsibility that I do feel is that seeing as I'm here in this body in this world I've got to act in this body in this world in an appropriate way and an appropriate way for me is trying to grow or get better and primarily for myself and secondarily to help other people". (Researcher; James, Affirmative Brief)
>
> I'm being realistic, she [his daughter] has caused me pain + that has to stop – today is the day to do that... It's almost the same with my girlfriend who had to take responsibility for her mental health + I will tell her I will leave if she doesn't take care. (Client; Simon, Affirmative Brief).

> With the support of the therapy, she was able to act out of her personal agency and was affirmed in validating and experiencing her ability to make choices for herself. (Adjudication panel; Sofia, Adjudication)

Sub-category 3.2.1.3.2.3: Increased awareness of impact on/of self in relationship. Some clients *increased their awareness of the impact that they had on others, and others had on them, in relationship* (improvers – variant):

> I reminded her [his therapist] that I just want to make sure they [his family] don't go into meeting with each other + with high expectations – start shouting and + get louder + then it's who shouts the loudest + competitive, as their stuff affects me. (Client; Simon, Affirmative Brief)

Sub-category 3.2.1.3.2.4: Figuring out who/how to trust. Finally, one improver expressed the importance for them in *figuring out who/how to trust:*

> Talking about trust in general has helped me a lot to figure out who/how to trust now as an adult and also to trust more in myself. (Client; James, Affirmative Brief)

Sub-category 3.2.1.3.3: Increased openness. The experience of increased openness (improvers – general; deteriorators – unique) was expressed in several ways: increased openness *to dealing with life as it comes* (improvers – typical); *to connecting with other people* (improvers – variant; deteriorators – unique); *to being in the moment* (improvers – variant); and *to taking risks / challenging self* (improvers – unique).

Sub-category 3.2.1.3.3.1: To dealing with life as it comes. It was typical for improvers to end therapy feeling more open *to dealing with life as it* comes and, in particular, that *some things were not in their* control (improvers – typical). For example:

> James is more able to deal with life as it comes: “It’s not always changing at the right rate or direction. I’m much more content now to just be myself I think that it doesn’t really matter what happens, I can deal with it.” (Researcher; James, Affirmative Brief)
>
> I talked about a disagreement with my friend. I realised what I want from life, and what I don’t want, especially from the people that call themselves friends. It made me realise that I have no control over events but I do have control over how I deal with them. (Client; Linda, Affirmative Brief)

Sub-category 3.2.1.3.3.2: To connecting with other people. Some clients discovered by the end of therapy an increased openness *to connecting with other people* (improvers – variant; deteriorators – unique), including being *more accepting of others’ differences* (improvers – unique; deteriorators – unique). For example:

> It was helpful for me to see/realize the many positive possible aspects of human interaction. The world seems like a nicer and less frightening place. (Client; James, Affirmative Brief)
>
> Maybe the relationship’s quite a big thing for me because, em, maybe not everyone wants to be with someone but we all have different ways that we look at our life and I really like sharing with someone, being close. (Client; Simon, Affirmative Brief)

Sub-category 3.2.1.3.3.3: To being in the moment. Two clients recognized an increased openness in themselves *to being in the moment* (improvers – variant). For example:

> [One of] Simon’s end of therapy [changes] was that: “I am more accepting of the past and living in the moment”. (Researcher; Simon, Affirmative Rebuttal)

Sub-category 3.2.1.3.3.4: To taking risks / challenging self. Finally, one client noted their increased openness *to taking risks / challenging self:*

> I am open to do more scary and unpredictable things. (Client; Julia, Affirmative Brief)

Sub-category 3.2.1.3.4: More positive / balanced / realistic perspective on life. I found evidence in the data of another change that related to self in the world: that by the end of therapy most improvers seemed to have developed a *more positive / balanced / realistic perspective on life* (improvers – typical). For example:

> I generally have a more positive outlook on everything. (Client; Julia, Affirmative Brief)

> This is helping me to be more realistic + move on + look forward to life rather than fear it. (Client; Simon, Affirmative Brief)

Sub-category 3.2.1.4: Positive change in personal circumstances. The final facilitative change experienced by clients that I noted in my analysis was a *positive change in personal circumstances* (improvers – typical; deteriorators – unique). Most clients not only experienced the positive impact of changes in their circumstances as a result of events that took place but there was also evidence of clients discovering that they had *increased resources* available to them (improvers typical; deteriorators – unique), either *within relationships* (improvers – variant; deteriorators – unique) or *within self* (improvers – unique). For example:

> Linda's outcome data indicates that she changed substantially not throughout therapy, but within the first two sessions. Paired with the information around Linda's extra-therapy events, this change would appear to be due to a significant change in Linda's life and work circumstances. (Researcher; Linda, Sceptic Brief)

> The course and offer of an interview is what has been missing for the last five years. I've been unemployed with major life events and poor health and now I can see more positivity in the future + if I can

> continue to work on my detailed plan + take action I will achieve all my goals of better health and back to employment that is meaningful. (Client; Simon, Affirmative Brief)
>
> I'd spoken about that with my counsellor, and that night my sister had come online, and I don't know how many sessions there had been... Maybe three or four... I don't know how many... … Ehh... Maybe, if I hadn't been speaking to my counsellor, then I wouldn't have been able to talk to my sister about it? (Client; Linda, Affirmative Brief)
>
> Coming up with an idea of how to reflect on my feelings by writing in a way in which could help to continue the progress that has already been made. I feel more encouraged and confident and hopeful as a result. (Client; James, Affirmative Brief)

Sub-category 3.2.2: Interfering Changes. My analysis identified four main *interfering changes* experienced by some clients, most commonly deteriorators, by the end of therapy (improvers – unique; deteriorators – general). These were: *deterioration / no change in key difficulties* (deteriorators – general), *disappointment with process* (deteriorators – general), *loss of hope* (deteriorators – general) and *negative change in personal circumstances* (improvers – unique; deteriorators – variant).

Sub-category 3.2.2.1: Deterioration / no change in key difficulties. Only deteriorators experienced *deterioration / no change in key difficulties* (deteriorators – general). For example:

> The first five longstanding problems on the list, therefore the most important to Luke, all deteriorated over the course of therapy. (Researcher; Luke, Affirmative Brief)
>
> At the end of therapy, three of these items still indicated deterioration from her post therapy rating, each of which were rated

> as long standing (6-10 years or more). (Researcher; Caitlin, Affirmative Brief)

It was noteworthy that, for some, this deterioration was experienced *later in therapy* (deteriorators – variant); for others it was indicated by a sense of being *no clearer at end of therapy* (deteriorators – variant). For example:

> Sofia's discomfort at living in the UK seems to become more prevalent towards the end of therapy. (Researcher; Sofia, Affirmative Brief)

> In his Working Alliance Inventory, Joseph records a decrease on items 1 and 2 which relate to his feelings about the potential for change and the role of the therapist in supporting him to look at his problems. For both of these items Joseph's score drops from 4 (very often) to 2 (sometimes) suggesting that, rather than feeling clearer about his long-standing problems, he feels at best no different and worst as though he is unable to make a change or that he is unable to be helped. (Researcher; Joseph, Affirmative Brief)

For one client, his *scores suggested he did better when not in therapy* (deteriorators – unique):

> The biggest reductions (therefore improvements to Luke) are between the intake and the first session, and between the 19th and 20th sessions… An ending occurred at the 19th session, and the gap between sessions 19 and 20, at 22 days, was the longest stretch of time Luke went without a session over the course of his therapy. This suggests that Luke did better when he wasn't in therapy. (Researcher; Luke, Affirmative Brief)

Sub-category 3.2.2.2: Disappointment with process. I found evidence that suggested growing *disappointment with the process* (deteriorators – general). In particular this disappointment was founded on two main experiences: *feeling worse*

(deteriorators – general) and a *perception that getting nothing from the process* (deteriorators – variant).

Sub-category 3.2.2.2.1: Feeling worse. Clients reported *feeling worse.* In particular they felt *more depressed* (deteriorators – general), which included being *less accepting of self* (deteriorators – variant) and *struggling to make peace with changes* (deteriorators – unique). For example:

> I came out feeling quite heavy about it and, er, and I had to go and meet a friend afterwards and it kind of affected my conversation with my friend because I was having to perk up for my friend when I've just come out feeling bit down about stuff. (Client; Joseph, Affirmative Brief)

> In his Mid-2 change interview (please note that this was, at the time, an end of therapy interview), Luke describes that he "felt more guilty about self", and states that it would have been "unlikely" to happen without therapy. (Researcher; Luke, Affirmative Brief)

Clients also described feeling *more vulnerable* (deteriorators – typical). This led some to becoming *less willing to engage emotionally* (deteriorators – variant) and to *feeling less supported* (deteriorators – unique):

> Furthermore, one of the changes that Sofia identified mid-therapy is both negative and concerning: "pulling my hair", indeed this is also something she contacts her therapist about in an email asking for help with this issue. This suggests that there was indeed deterioration in therapy which led to her harming herself. (Researcher; Sofia, Affirmative Rebuttal)

> Luke was very emotionally shut down and defeated even towards the end of therapy, writing in an email "What's the point of getting annoyed at something if later on it loses all the meaning… At first it hurts but then you forget that it's there at all and by the time you know it you let people walk all over you and use you and you just go

with it because there's nowhere to run." (Researcher; Luke, Sceptic Rebuttal)

Sub-category 3.2.2.2.2: Perception that getting nothing from the process. A second feature of these clients' growing disappointment with therapy is demonstrated in their descriptions of *getting nothing from the process* (deteriorators – variant). For example:

> It's just sometimes, it depends on the mood I'm in, if I come here and I've nothing to talk about I feel like I'm wasting my time and their time. (Client; Joseph, Affirmative Brief)

> Another notable hindering aspect shared in this interview was that she had been: "…unpacking stuff but no resolution yet." (Researcher; Caitlin, Affirmative Brief)

Sub-category 3.2.2.3: Loss of hope. There was evidence that *loss of hope* during the period of therapy affected deteriorators in particular (deteriorators – general), reinforced by a growing sense that there would be *no resolution* (deteriorators – typical):

> At the time when Joseph ceased therapy he did not engage in any ending or completion of therapy and did not attend for his end of therapy interviews, suggesting that he felt a further change of therapist was not helpful to him and that he felt that therapy was not helpful to him. (Researcher; Joseph, Affirmative Brief)

> The nature of Caitlin's ending suggests that she was not ready for the intensity of her emotions if she were to allow things to "come to a head" without the comfort of feeling like there could be a resolution (Researcher; Caitlin, Affirmative Brief)

Sub-category 3.2.2.4: Negative change in personal circumstances. Some clients experienced *negative change in their personal circumstances* during therapy (improvers – unique; deteriorators – variant). For example:

> The therapist considers that Simon was a bit down regarding the outcome of his welfare appeal, adding that he said that it felt like a bit of an interrogation which was a difficult process to get through. (Researcher; Simon, Sceptic Brief)
>
> In addition to noting that she had stopped her medication at Mid-3, Caitlin also experienced a significant life-event within this same time period involving the death of her father. In her communication of this to her therapist via email, she does express that his death causes her to experience physical symptoms of anxiety which prompt her to enquire about a sooner counselling appointment date. (Researcher; Caitlin, Sceptic Brief)

Category 3.3: Potential impact of research. The HSCED process evaluates the potential impact that being a research participant may have had either on the outcome of the therapy itself or on the reliability of the data collected as part of the process. Therefore, it was possible to include this data within my analysis and to highlight any apparent similarities or differences between the two groups that appeared to exist. The data that I found in the HSCED material organized itself into three sub-categories: *data may not be accurate* (improvers – general; deteriorators – general), *increased self-worth* (improvers – variant; deteriorators – unique), and *the PQ provided motivation and focus* (improvers – unique; deteriorators – unique).

Sub-category 3.3.1: Data may not be accurate. There was evidence that the data collected from clients in both groups *may not be accurate* (improvers – general; deteriorators – general). This argument was supported by a variety of plausible reasons: that the client *may have wanted to please the therapist / researcher* (improvers – typical; deteriorators – typical); that it was *hard to capture the experience* using the research instruments (improvers – typical; deteriorators – variant) – for example, that it was *hard to shift between the forms and the session* (improvers – unique; deteriorators – unique), or because the client *sometimes felt tired or unwell* (improvers – unique; deteriorators – unique), or that *no adjustments were made* for a client with autistic processes (improvers – unique); that some

clients felt *self-conscious* (improvers – unique; deteriorators – variant), and because the *amount of paperwork was hindering* (deteriorators – unique). For example:

> [At her six months follow up interview with a different researcher] changes were attributed significantly less to therapy, with two out of four rated somewhat likely without therapy, one rated somewhat unlikely, and only one rated unlikely without therapy […] Whilst this may be due to a longer period of retrospection, it could also support the argument that Linda was more likely to attribute post and mid therapy changes to the therapy with her initial researcher. (Researcher; Linda, Sceptic Brief)

> It also a partial part of his experience that there were some occasions in which he felt that doing some of the forms were too much for him for various different reasons. (Researcher; Simon, Sceptic Brief)

> "The end of the interview questionnaire or the end of session questionnaires. I remember sometimes thinking that they were hard to fill out because we just had an hour of intense talking about feelings and I couldn't even remember what we started talking about sometimes." Here Julia reports how after the end of session she wouldn't remember most of her experiences in session and it can be argued that her positive score might be artificially inflated resulting in unreliable data. (Researcher; Julia, Sceptic Rebuttal)

> Luke did not mention this unprompted, but when asked in the change interviews he responded, in three of the five interviews, that he found the amount of paperwork hindering. Although this in and of itself may not have been a large factor in Luke's deterioration, it might have contributed to his overall feeling that he was wasting his time. (Researcher; Luke, Affirmative Brief)

Sub-category 3.3.2: Increased self-worth. Some clients experienced an *increased self-worth* as a result of their participation in the research activities

(improvers – variant; deteriorators – unique). This was for a two main reasons: because they were *helping others* (improvers – variant; deteriorators – unique) and *helping self* (improvers – variant; deteriorators – unique):

> Being part of the research study gave the client a sense of value and relevance he might not have otherwise felt. The client frequently shared that he felt a strong sense of importance, value and giving back as a participant in the research study. (Researcher; James, Sceptic Brief)

> I'm quite happy that I had these sessions free of charge for all this time. It's made a massive difference to me. (Client; Simon, Sceptic Brief)

Sub-category 3.3.3: PQ provided motivation and focus. Two clients (improvers – unique; deteriorators – unique) described the *motivation and focus* that they experienced from using the weekly outcome measure, the Personal Questionnaire, as part of their therapeutic process. For example:

> Regarding her use of monitoring tools, she expresses the personal benefit she experienced in being able to view her feelings from previous weeks in retrospect. She also noted in her end of therapy interview that completing a weekly PQ gave her a way to "pay attention" to her problems. (Researcher; Caitlin, Sceptic Brief)

In the End: Summary. As before, this section summarises the general and typical findings within the domain for each group. These are presented in Table 6.5.

The analysis identified distinct differences between the two groups in the *ending* of their therapeutic experience. The data suggests that only improvers experienced *facilitative aspects of ending*: for example, they *felt ready to end*, they believed that the *sessions helped a lot* and that they had *made great progress* and there was a *mutual decision to end*. In stark contrast, deteriorators experienced an *incomplete therapeutic process* in which the decision to end was initiated by their

therapist's decision to leave (permanently or as an extended break), leading to the *client making the decision not to continue*.

Table 6.5. Domain 3: In the End

Category	Improvers	Deteriorators
Ending of therapy	**Facilitative aspects of ending: Feeling ready to end; Sessions helped a lot -> made good progress; *Mutual decision to end***	**Interfering aspects of ending: Incomplete therapeutic process -> C*lient decision not to continue following therapist decision to leave / take extended break***
Changes	Facilitative changes: Improvement in key difficulties that wanted to address in therapy -> ***during early stages in therapy*** Relationship with self -> increased self-awareness / self-understanding -> more aware of needs and their impact; **increased self-acceptance -> more self-compassion*, increased appreciation of self;* feeling complete / more integrated / back on track -> *increased self-control/ stillness/ calm*** Self in the world -> **more empowered/ motivated -> *increased self-confidence / self-trust / self-belief;*** *new understanding of self in relationship ->* ***less fearful of abandonment & rejection;* increased openness-> *to dealing with life as it comes -> some things not in control; more positive/ balanced/ realistic perspective on life***	Facilitative changes: Improvement in key difficulties that wanted to address in therapy; Relationship with self -> increased self-awareness/ self-understanding -> *more aware of needs & their impact;* Self in the world **Interfering changes: Deterioration/ no change in key difficulties; Disappointment with process -> feeling worse -> more depressed/ distressed; *more vulnerable;* Loss of hope -> *no resolution***
Potential impact of research	Data may not be accurate -> *may have wanted to please therapist / researcher*; *hard to capture experience*	Data may not be accurate -> *may have wanted to please therapist / researcher*

Notes. **Bold** text = difference of at least two frequency counts; non-italicized text = general frequency; italicized text = typical frequency.

There were similarities and differences in the nature of the *change* experienced by clients in the two groups. There was evidence that both sets of clients benefited from *facilitative changes,* including *improvement in key difficulties that they had wanted to address in therapy* but that improvers noticed *improvement during early stages in therapy*.

Changes in their *relationship with self* was also noted for clients in both groups finding, in particular, that *increased self-awareness / self-understanding*, including being *more aware of their needs and their impact* on them was experienced by improvers and deteriorators alike. However, the changes in this area noted for improvers also included *increased self-acceptance*, with *more self-compassion* and *appreciation of self*, and *feeling complete / more integrated / back on track*, including *increased self-control / stillness / calm*.

There was evidence of change relating to *self in the world* amongst clients in both groups. Improvers appeared to have moved further into this process than deteriorators, with evidence that they were *more empowered / motivated*, with *increased self-confidence / self-trust / self-belief,* had a *new understanding of self in relationship,* noticing that they felt *less fearful of rejection and abandonment, increased openness*, especially to *dealing with life as it comes* because *some things are not in their control*, and a *more positive / balanced / realistic perspective on life*.

The final area of *facilitative change* experienced primarily by improvers was *positive change in their personal circumstances*, including *increased resources*.

Evidence of more *interfering changes* was largely limited to the data gathered from the deteriorators. I noted that all deteriorators experienced *deterioration or no change in some key difficulties*; that they developed *disappointment with the process*, fueled by *feeling worse*, in particular *more depressed and distressed,* and *more vulnerable*, resulting in a *loss of hope* and sense that there would be *no resolution*.

Finally, I reviewed the *potential impact* that participation in the *research* may have had on clients in both groups. I found evidence that there was an equal likelihood that the *data collected may not be accurate*. Typically, for both groups, this might be because the client *may have wanted to please the therapist or researcher*. In addition, improvers commented that at times it was *hard to capture their experience* in the forms provided.

Chapter Summary

In this chapter I adopted a two-stage mixed method approach to investigate if a change in scores recorded on the Strathclyde Inventory accurately reflects a participant's therapeutic outcome, and to identify what features of the therapeutic experience may explain client improvement or deterioration at the end of therapy as measured by the SI.

First, I supervised the creation of a systematic case study series focused on eight participants from my Study 2 dataset. My co-researchers used the critical-reflective, quasi-judicial HSCED method (Elliott, 2014, 2015) to evaluate the outcome of their therapy based on the full rich case record gathered throughout their counselling process at the research clinic: a wide range of qualitative and quantitative data collected from both the client and the therapist. This first stage culminated in consensual decisions made by adjudication panels that appeared to match the relative degree of change indicated by the SI standardised residual scores recorded by participants. This result indicates that there was an acceptable degree of convergence between pre-post change in SI scores and the overall assessment of outcome captured by the HSCED process, according to this small study of eight participants whose pre-post change in SI scores was capable of being defined as reliable change. Further testing is required to evaluate the outcomes of clients whose SI scores indicated no change by the end of therapy.

The adjudication panels also confirmed that sufficient evidence was provided to support the affirmative proposition that each individual client's experience of therapy had some degree of impact on the change that occurred,

ranging from *slightly-moderately* in the case of Sofia, a deteriorator, to *substantially* in the cases of Julia and Simon, both improvers. The panels provided details of the arguments and evidence drawn from the affirmative and sceptic cases, including proposed mediators and moderators that informed their decisions.

This is the first study in which the HSCED method has been offered as a contribution to the assessment of the validity of an outcome measure used in counselling and psychotherapy research. I believe that this study has demonstrated the potential for integrating this approach with more conventional measure development processes. The HSCED method provides a model in which the one-dimensional snapshot of outcome created by scores collected on an instrument can be robustly evaluated from the perspectives of two narratives constructed using multi-dimensional data collected not only from the client but also others involved in their process, most typically the therapist. It provides a genuine link between more abstract forms of research (the measure) and the realities of practice (the client's lived experience in therapy), ultimately evaluated by an adjudication panel that aims to reach a consensus in their understanding of what has taken place. There is a clear and valuable opportunity for assessing the validity of an outcome measure – is it measuring the type of change that it proposes to measure? - within this process.

This study benefited from the opportunity to test the HSCED model with a series of eight case studies. Being able to compare the panels' decisions enabled me to get a sense of the degree to which they matched, in order and distribution, with the participants' standardised residual scores. Including both *good outcome* (i.e. improvement) and *poor outcome* (i.e. deterioration) cases and creating two groups of researchers, who focused on one particular type of outcome and then performed adjudication duties for the other group, enabled me to test the process of adjudication panels making decisions by consensus and also consider the potential influence of panel perspectives on their decisions, noting in particular a potentially 'softer' evaluation of deterioration by the panels who had been working on developing the improvers' case studies. Carrying out the meta-synthesis enabled me to gather together all eight cases and gain a clearer understanding of what can

be learned about these clients' therapeutic outcomes, including the common and differing features in their therapy that may have led to these outcomes, and as a result conduct a further assessment of the validity of change in scores on the SI as a measure of a client's therapeutic outcome.

To What Extent does the Meta-synthesis of the HSCED Results Validate Change in SI Scores?

My meta-synthesis of the HSCED materials enabled me to identify very distinctive patterns in the data: similarities in the features of the therapeutic experience shared by improvers and deteriorators respectively, and differences between the two groups that may explain the improvement or deterioration in their SI scores at the end of therapy. These findings confirmed and validated the pre-post change in scores recorded by the two groups of participants. However, more than that, the findings in the meta-synthesis offer a tentative explanation for the different capacities for congruent functioning identified by the post-therapy scores captured by the SI at the end of therapy.

As the third domain of my meta-synthesis describes, improvers felt ready to end therapy and the decision to end was by mutual agreement with their therapist. They left therapy feeling more self-aware, more self-accepting and self-compassionate and increased openness to being themselves when relating with others out in the world. This synthesised description of the experiences of these four clients is a clear match for the theoretical construct of congruent functioning and also the SI developmental pathway, based on the Rasch hierarchy of items, proposed in Chapter 4. I will discuss the contribution these findings make to knowledge in the next chapter.

In contrast, the end of therapy for deteriorators typically occurred at a time not of their choosing but instead as a result of the needs of their therapists. Although these clients could have continued in therapy, they decided that they would not, more than likely because they were feeling disappointed, depressed, distressed and vulnerable, and had lost hope in the process. They also described and provided evidence of becoming more self-aware over the course of therapy, in

particular more aware of their needs and the impact that these have on them. Considering this synthesised description of their therapeutic outcome, it seems plausible that these clients ended therapy at a point at which, paradoxically, they were capable of greater congruent functioning (i.e. self-awareness) than when they started but were unable to move through this process sufficiently to experience the self-acceptance and self-compassion that is essential in order to ease the pain of this experience and was achieved by the four improvers in this study. A similar distinction was identified by McElvaney & Timulak (2013, p.251) in their comparison of the experience in therapy of 'good' and 'poor' outcome clients.

So, what does this mean for the validity of their SI scores and the change in scores reported? I propose that the deterioration in their scores from the beginning to end of therapy represents the process of becoming more self-aware, the starting point for increasing congruent functioning. This evidence supports my hypothesis that, at least for some clients, scores on the SI will decrease as they become more accurately representative of the individual's awareness of their current experience. In contrast to the patterns indicated by six of the seven participants highlighted in Chapter 5, the process for these clients occurred over the whole course of therapy and they did not experience a turning point that enabled them to grow further. As a result, I propose that the deterioration indicated in the pre-post change in SI scores can be understood as representing an incomplete therapeutic process, rather than the assumption that there has been a decrease in their capacity for congruent functioning. Their process in therapy was stymied for a variety of reasons, discussed below, and the circumstances that led to the end of therapy meant that each client ended therapy without gaining the outcome that they had hoped.

To What Extent are Therapeutic Processes Associated with Person-Centred Theory found in the HSCEDs Causally Related to Change in SI Scores?

Reflecting on my findings, I believe that the key features in the therapeutic experience that distinguished between clients whose SI scores improved, and those whose SI scores deteriorated, were: (1) the client's relative readiness for therapy and the impact that unresolved doubts had on their experience of the subsequent

therapeutic process (i.e. discomfort, deference, and difficulty integrating therapy and life); (2) the impact on the therapist of their client's struggles within the process (i.e. their own frustration and doubt); and (3) the impact of disruption and delay on clients and the consistency and development of their therapeutic process.

First, the meta-synthesis found differences between improvers and deteriorators in relation to their readiness to begin therapy. Improvers tended to demonstrate more of the characteristics that indicated readiness: motivation, determination, openness to experience, the ability to be reflective and preparedness for, and, in the process. Although the deteriorators were motivated to participate in therapy, they also carried doubts about the process or the potential for doubt, for example because they expected to feel more vulnerable or because they had the 'wrong' expectations of this form of therapy, such as receiving advice from their counsellor, similar to the *initial misgivings about psychotherapy* reported by MacFarlane, Anderson & McClintock (2015). It seems to me that the impact of these doubts or potential doubts may have been somewhat disregarded or underestimated by their therapists, a finding shared by Werbart, von Below, Engqvist and Lind (2018) and Werbart, Annevall & Hillblom (2019). Certainly in this study, it appears as if these doubts and potential doubts were subsequently played out within the therapeutic process: deteriorators experienced discomfort in the process, they found it painful and difficult to open up, felt exposed and struggled to find direction; they deferred to their therapist, were concerned about upsetting them and felt that they had to do therapy differently; some demonstrated a lack of engagement while at least one described themselves as coasting. It seems likely that all of these difficulties contributed to the typical experience shared by deteriorators in which they struggled to integrate therapy and life, in sharp contrast to the more positive experiences of the improvers.

The next part of the puzzle, I suggest, was the impact that these difficulties experienced by their clients in the process had on their therapists and their response. The meta-synthesis showed that there was strong evidence of the deteriorators' therapists validating and affirming their client's experience, thereby

promoting the client's awareness of themselves within their situation and initiating an increase in their capacity for congruent functioning, but also that at times the therapists worked outside their client's frame of reference and could act in a controlling or directive manner: unhelpful therapist behaviours also identified by Curran et al. (2019). From the perspective of person-centred theory, this suggests that these counsellors were experiencing doubt and frustration with themselves, their clients and the process and, therefore, unable to maintain the six conditions required for therapeutic change (Rogers, 1957/1992) to a sufficient degree. In particular, I propose, their capacity for experiencing unconditional positive regard (UPR) and empathy towards their clients may have been impaired and therefore their ability to maintain psychological contact with their clients, and for their clients to perceive their empathy and UPR, may have been affected. Under such circumstances, it is no wonder that the clients and therapists struggled to find a way to work through their mutual doubts, potentially unspoken, that were impacting on their work together.

These difficulties may be understood in the context that these trainee therapists had perhaps not yet expanded their own capacity for congruent functioning, developed their confidence to use *metacommunication* (e.g. Mearns & Thorne, 2013, pp.110-112; Rennie, 1998, pp.89-101) to explore *the unspoken relationship* (e.g. Mearns, 2003, pp.64-74) or indeed encountered the research-informed work of Rennie, who contributed concepts such as client deference, therapist self-disclosure and metacommunication to the development of person-centred theory (Levitt, Lu, Pomerville & Francisco, 2015). Nevertheless, these difficulties within the therapeutic relationship are not ones that are exclusive to trainee therapists, replicating the findings of several recent studies (e.g. Curran et al., 2019; Hardy et al., 2019; Werbart et al., 2018; Werbart et al., 2019) in which experienced therapists' negative feelings and reactions to their clients and the absent or unsuccessful discussion between therapists and clients about concerns within their relationship, were perceived as contributing to lack of progress in the therapy.

These apparent difficulties in the therapeutic relationship for deteriorators and their counsellors seem to have been further exacerbated by disruption and delay in the process. Although two improvers also faced the challenge of changing therapist during their involvement with the research clinic, they used the experience as a way of assessing their growth (i.e. how well they handled the change of therapist). It appeared to have only beneficial impact on them. In contrast, while some deteriorators were optimistic about the change and there was evidence that their new therapists may have been a better fit for them, the change typically occurred at an unfortunate time for the client for a variety of reasons. It seems highly likely that this unsettling change, which required clients who had come to therapy reporting difficulties in their relationships with others to begin another new relationship, combined with an inconsistency in the frequency of sessions for some, would have added further discomfort for deteriorators in the process of developing a sufficiently therapeutic relationship that would enable them to engage more fully in the process. It is no wonder that, when faced with the prospect of changing therapist again or to have an extended break initiated by the counsellor, these clients made the decision to end their participation in the process.

The findings of this meta-synthesis provide important insight into a growing area of the literature: negative experiences in counselling and psychotherapy. As the study indicates, both the client and therapist can struggle within the process and it is inevitable that this will impact on the client's therapeutic outcome. As Hardy et al. (2019, p.411-412) have described, it is rarely a single contributory factor that causes a negative experience but rather client and therapist becoming "stuck in negative interactional pattern from which change [becomes] impossible". I will discuss implications for practice in the next chapter.

Chapter 7: Discussion

This thesis, an investigation of the Strathclyde Inventory (SI) as a measure of outcome in person-centred therapy, was developed to answer two overarching questions:

1. Are scores on the Strathclyde Inventory a valid measure of congruent functioning when used with a UK-based clinical population?
2. Do scores on the Strathclyde Inventory change over the course of therapy in a way that is consistent with person-centred theory?

In this discussion chapter, I will collate the key results from my three studies and assess the answers that they provide to these questions. In addition, I will identify the implications of my findings for theory, research and practice, the original contribution to knowledge that my thesis offers, and conclude with some reflections on my own learning through the process.

Summary of Findings

Are Scores on the Strathclyde Inventory a Valid Measure of Congruent Functioning when used with a UK-based Clinical Population?

As I outlined in Chapter 3, validity is a fundamental requirement for any instrument designed for use as a measurement tool in counselling. In this thesis I followed contemporary guidance by approaching validity as a unified concept in which the potential validity of an instrument relates to the strength of evidence arising from the data it generates, including evidence of reliability/precision and issues relating to fairness, and demonstrates its ability to capture the theoretical construct it is designed to measure, its hypothesised process of change, and relevant relationships with other variables.

Across my three studies, I have amassed a range of evidence that supports the validity of using the SI to measure congruent functioning, and change in congruent functioning, with data collected from clients in therapy at the UK-based Strathclyde Counselling and Psychotherapy Research Clinic (‘the research clinic’).

In my first study (Chapter 4), I found evidence of excellent internal consistency and precision/reliability, a well-functioning 5-category rating scale, good fit with a unidimensional model thereby resolving a long-standing question about the SI structure, and the ability to distinguish meaningful levels or degrees of congruent functioning, inferring a potential developmental pathway. I produced a 12-item version that maintained these characteristics, according to the scores in my dataset, in order to reduce the administrative burden and increase the potential fairness of its use for, and with, future clients completing the SI alongside their counselling process.

In my second study (Chapter 5), I found evidence that the SI is sensitive to change in scores when data is collected from clients at the beginning, during, and at the end of therapy. My analyses demonstrated that the extent of pre-post change across therapy experienced by the participants in my dataset, as captured by the SI, was equivalent to other outcome measures used with clinical populations, both in this study and others, whether analysed using effect size or clinical significance methods. When looking more closely at patterns of change in scores during the process of therapy, I identified: (a) a potential relationship between reliable change in scores at data collection time points and the decision to end therapy that appears to provide evidence of the *good enough level* model (Barkham et al., 2006): and (b) a cubic polynomial trend for improvement in congruent functioning, with significant change occurring at both early and late stages of therapy, consistent with Owen et al.'s (2015) *early and late change* trajectory. These results demonstrated that the SI is capable of capturing change in scores and that the patterns of change being detected across therapy are consistent with findings generated in comparable studies of other validated instruments.

In my third study (Chapter 6), I introduced an innovative mixed method approach to examine the validity of change in SI scores in practice through the example of eight individual clients: four whose change in scores indicated reliable improvement across therapy, and four whose change in scores suggested reliable deterioration. By conducting a meta-synthesis of the data produced from these

systematic case studies, I found evidence of distinctive similarities and differences in the therapeutic experiences of clients who improved or deteriorated according to pre-post change in their SI scores. These findings offered a consistent and credible explanation of the different outcomes experienced by these eight clients, according to their scores on the SI, thereby providing evidence that change in scores collected on the instrument across therapy can be considered an accurate reflection of the client's experience at the end of therapy.

Overall, I believe that the weight of evidence generated by this series of studies provides a clear basis for accepting the validity of using and interpreting scores on the Strathclyde Inventory as a valid measure of outcome in person-centred therapy. There are, of course, potential limitations to the scope of this conclusion, which will be highlighted later in this chapter.

Next, as an essential aspect of validity is the degree to which the theoretical construct that the instrument intends to measure is consistent with the theory itself, I will identify the main findings that relate to the concept of congruent functioning and explore how these support and develop current understandings in person-centred theory, beginning with the evidence generated in my series of studies that contribute an answer to the second overarching question that guided this thesis.

Do Scores on the Strathclyde Inventory Change over the course of Therapy in a Way that is Consistent with Person-Centred Theory?

The results of my second study confirmed that the shape of change in clients' scores collected on the SI over the course of therapy demonstrated a high degree of variation. The recognition of the idiosyncrasy in each client's experience in therapy is at the heart of person-centred therapy (e.g. Keys, 2003). Nevertheless, my analyses also demonstrated that patterns across the group data were detectable, offering the potential for interpretation and comparison with specific aspects of person-centred theory.

The findings in this study provided preliminary evidence of: (a) a relationship between increased congruent functioning (as measured by the SI) and reduction of distress (as measured by the non-risk scale of the CORE-OM); (b) an apparent response shift (McLeod, 2001) that suggested a change in clients' perception of the meaning and relevance of specific items between the beginning and end of therapy, inferring increased congruent functioning in the form of self-awareness developed over the course of therapy; and (c) examples of decreasing scores between data collection points over the course of therapy that may indicate a process of increasing congruent functioning in which scores become a more accurate representation of the individual's perception of self within their current situation. The relationship between these specific patterns of change, and aspects of person-centred theory, were outlined in Chapter 5.

The results of my third study distinguished key features of the experience in therapy for clients whose SI scores improved by the end of therapy, and those whose scores deteriorated, appearing to explain why their scores changed in the direction that they did. In summary, these were: (a) the client's relative readiness for therapy and the impact that unresolved doubts had on their therapeutic process; (b) the potential impact on the therapist of their client's struggles within the process and their ability to respond; and (c) the potential impact of disruption, delay and inconsistency on the development of the therapeutic relationship. I related these findings to aspects of person-centred theory in Chapter 6. Finally, a comparison of the changes experienced by the two groups of clients at the end of therapy, gleaned from an evaluation of the full range of quantitative and qualitative data collected from them, reflected and elaborated the developmental pathway for congruent functioning identified in my first study. I will discuss this finding in more detail in the next section.

Taken together, I propose that these results provide evidence that scores on the SI, with this particular client group, do change over the course of therapy in a way that is consistent with person-centred theory. The implications for theory arising from this conclusion are discussed in the next section.

Implications for Theory

I have identified the main implications for theory arising from this thesis as: (a) a developmental pathway for congruent functioning; (b) one pathway, many routes; (c) deterioration in congruent functioning as the outcome of therapy indicating an incomplete process; and (d) self-acceptance as the pivot point in congruent functioning. I will discuss each of these areas below.

A Developmental Pathway for Congruent Functioning

A major development for theory was the identification of a hierarchical relationship between SI items based on a Rasch calculation of probability arising from patterns in scores collected from participants at all stages of the therapeutic process. This hierarchy of items offered a deconstruction of the theoretical construct, congruent functioning, as measured by the SI. Based on my interpretation of the content and grouping of these items within the hierarchy, this indicated that self-awareness was the least difficult form of congruent functioning, according to scores on the instrument, followed by self-trust, self-acceptance, openness to self, and then, most difficult of all, openness to others.

As this finding was based on one-off observations provided by participants, a potential developmental pathway – a process of increasing congruent functioning – could not be assumed. However, the findings of my third study provided further elaboration of this proposed model. Specifically, the facilitative changes identified within the *In the End* domain demonstrated that, while both groups of clients developed their relationship with self, experiencing an increase in their self-awareness and self-understanding by the end of therapy, only those clients whose scores identified them as improvers also increased in self-acceptance, reporting that they felt more complete, integrated and 'back on track'. In addition, improvers reported changes that related to their sense of self in the world: feeling more motivated and empowered (arising from increased self-confidence, self-trust and self-belief); new understanding of themselves in relationship, including becoming less fearful of abandonment and rejection; and an increased openness to dealing with life as it comes, and having a more positive, balanced and realistic perspective

on life. In other words, those clients whose scores identified them as deteriorators reported changes at the end of therapy consistent with congruent functioning at a relatively early stage of development, according to the model. In contrast, the range and type of change reported by improvers was consistent with the fuller experience of congruent functioning depicted by the model, supporting the general trend in development proposed, while overlapping, intertwining and elaborating the descriptions provided by the items themselves.

The developmental pathway depicted is consistent with both Rogers' *19 propositions* (Rogers, 1951) and his *process conception of therapy* model (Rogers, 1961; Walker et al., 1960) and emphasises the relational nature of full congruent functioning. Indeed, it demonstrates that we can only become truly open to others (i.e. lower our guard, remove our mask) when we trust, accept and are able to be open to ourselves. Each layer of development expands our capacity for congruent functioning. As Rogers (1951, p.520) outlined in his 18th proposition:

> When the individual perceives and accepts into one consistent and integrated system all his sensory and visceral experiences, then he is necessarily more understanding of others and is more accepting of others as separate individuals.

This finding is also consistent with the more recent work of Stevens (2017) who proposed, as a result of his investigation of the role of authenticity as a mediator in the relationship between attachment style and affective functioning, that "if individuals cannot be genuine with themselves, then genuine behaviors and genuine relationships will be hard to establish" (p.408).

One Pathway, Many Routes

My first study demonstrated that data collected on the SI fit a unidimensional model, indicating that congruent functioning can, as Rogers (1961) proposed, be understood as a continuum of experience. Furthermore, as the analysis of the shape of change in scores over the course of therapy in my second study confirmed, a wide variety of potential routes may be taken within the

development pathway. As discussed above, I was able to identify preliminary evidence of patterns in the data that may provide signposts for a range of potential routes; here, I want to discuss one with specific theoretical significance in which I had been curious from the beginning.

When a decrease in scores implies an increase in functioning. I embarked on this research with a hypothesis that, for some, scores on the SI would decrease as they become more aware of the impact on them of the difficulties that they were experiencing in their life. This is a phenomenon recognised in practice in which some clients find themselves feeling worse as they begin to open up suppressed feelings and experiences during the counselling process and become more aware of their pain. In theoretical terms, their functioning was incongruent to the extent that they were unaware of their feelings and the impact of their experience. I had expected that this would be reflected in a decrease in scores on the SI, suggesting deterioration, early in the therapeutic process. In the course of my second study, I identified that reliable deterioration in scores between data collection points was not common (15 clients in total; 2-5% of clients at each time point), and that the largest incidence occurred between 10 and 20 sessions of counselling. This was an interesting finding as it demonstrated that this phenomenon may not necessarily occur only at the start of therapy. Instead, it implies that some clients may require a considerable time in therapy before becoming ready to make contact with their self-experiences.

A related discovery was that a reliable deterioration in scores between data collection points did not necessarily predict a deterioration in scores overall. Of the seven clients who continued in counselling beyond the point at which they registered a reliable decrease in scores on the SI, only one continued this trajectory to the end of their time in therapy. Three of the others ended therapy with an overall reliable improvement in scores, while the post-therapy scores of the other three were not significantly different to those recorded at pre-therapy, therefore obscuring the fluctuation in scores that occurred in the time between. This finding provides preliminary evidence that experiencing a decrease in scores on the SI

during the course of therapy is part of the process for some clients, reflecting Owen et al.'s (2015) *worse before better* trajectory. I propose that this process, identified as a response shift in the literature, is evidence of increasing congruent functioning during the therapeutic process.

Deterioration as Outcome: An Incomplete Therapeutic Process

However, some clients ended therapy when their pre-post SI scores indicated that a reliable deterioration had occurred. The findings from my third study enabled some investigation into what was happening for four of those clients in that moment and to contrast that with the experience of four improvers. I found a stark difference between the two groups at the end of counselling. While improvers felt ready to finish and had mutually agreed the decision with their counsellors, deteriorators disengaged from the process, typically in response to their therapist's decision to leave or take a break. They described feeling disappointed and distressed and, in this context, it makes sense that they decided not to continue with therapy having lost hope that the process would enable them to resolve their difficulties.

This evidence of an *incomplete therapeutic process* in 'poor' outcome cases was also identified by Watson, Goldman and Greenberg (2007) when investigating the experience of clients with depression who took part in emotion-focused therapy. Specifically focused on the limits of a short-term therapeutic process, Watson et al. proposed that their clients came into therapy at different stages of development and, therefore, some needed more time to develop their capacity to use the therapy as productively as others. However, similar to the *interfering therapist processes* found in my third study, they also noted that the therapists in these cases were aware that "things were not meshing with their clients" (Watson et al., 2007, p.201), and felt helpless and uncertain about how to address the relational and process difficulties that ultimately led to an unsatisfactory outcome for their clients.

The findings of my third study offer some potential guidance for therapists about the interfering aspects in the therapeutic process that appear to have inhibited the experience for the clients in this study and form the basis of the implications for practice that I discuss later in this chapter. Meanwhile, my findings have also provided an indication of where in the process of developing congruent functioning those clients seemed to become 'stuck', discussed above, and what might have occurred if their experience of therapy had enabled them to stay with the process, discussed next.

The Pivot Point in Congruent Functioning: Self-acceptance

Identifying the different degrees of congruent functioning attained by the group of improvers compared to the group of deteriorators at the end of therapy has provided empirical data that supports the significance of self-acceptance - or unconditional positive self-regard (Bozarth, 2001a) - as a pivot point in the therapeutic process. The results of my third study indicated that the group of clients whose scores had deteriorated by the end of therapy had increased their self-awareness but, in general, were unable to move forward into the development of self-acceptance. The conclusion proposed by this study is that this inability to progress was due to the difficulties that they encountered in the therapeutic process, which contrasted starkly with the experiences reported by the group of improvers. Indeed, if – as Bozarth (2001b) has argued – the therapist's unconditional positive regard is the "curative variable in [person]-centred therapy" (p.5), then these results provide evidence that the development of the client's reciprocal ability to develop self-acceptance is the curative experience – the mediator or causal mechanism - that enables further growth in congruent functioning.

The process of developing self-acceptance corresponds with Rogers' (1959) theory that reducing self-discrepancy is a mechanism for change. This finding supports the results of Watson et al.'s (2014) research into change in self-discrepancy, anxiety and depression over the course of therapy.

This finding also links with Neff's work on self-compassion. Neff et al. (2018, p.627) described self-compassion as "the balance between increased positive and negative self-responding to personal struggle", noting that self-compassion involves "being kinder and more supportive toward oneself and less harshly judgmental"; in other words, self-acceptance. The programme of research conducted by Neff and colleagues (e.g. Neff, 2003; Neff et al., 2017; Neff et al., 2018) has consistently demonstrated a relationship between self-compassion and psychological well-being.

Furthermore, my proposal that self-acceptance is a mediator for increasing congruent functioning complements the work of Pascual-Leone and colleagues who have conducted a significant programme of research into *emotional transformation* as a causal mechanism of change in psychotherapy (e.g. Pascual-Leone, 2018). This research has produced a sequential model of emotional processing that depicts a developmental movement from *global distress* to *acceptance and agency*, by working through maladaptive emotions (e.g. rejecting anger and shame/fear) in a way that enables negative self-evaluations and unmet existential needs to be expressed, producing a "categorically new experience [that leads] the client to a sense of 'Self as deserving' and mobilizes her or him to directly address unmet needs", and resulting in a sense of closure or resolution (Pascual-Leone, 2018, p.168). It is my suggestion that these two models are complementary, intersecting at the pivot point of self-acceptance, with the emotional transformation model, the core change concept in emotion-focused therapy (Elliott et al. 2004), providing more detail about the process of developing self-awareness, and the congruent functioning model illustrating the development of agency through growing openness to self and others.

Thus, the evidence identified in my thesis for self-acceptance as the pivot point for developing congruent functioning within the therapeutic process resonates not only with person-centred theory but also emotion-focused therapy and broader contemporary research in the fields of self-compassion and emotional transformation. My recommendations for research that can further investigate this

phenomenon are discussed in the next section along with other implications for research arising from this thesis.

Implications for Research

There are several limitations to my thesis that I will outline first, before presenting my recommendations for future research studies that build on my findings.

Limitations

The main limitations implicit in this three-part study relate to: (a) the use of archival data; (b) generalisability; (c) piloting adapted methods; and (d) researcher allegiance.

Use of archival data. The protocols for data collection in place at the research clinic are detailed and comprehensive, ensuring that a diverse range of potential studies can be conducted using data from the archive. While some of the common disadvantages of using archival data - e.g. obtaining permission to access, identifying an appropriate dataset (Jones, 2010) - were reduced because of my existing involvement with the research clinic and familiarity with the archive, I did notice limitations related to my choice to use secondary, rather than primary, data. This mainly took the form of missing data: actual missing data, e.g. a copy of the Strathclyde Inventory was not available at every data collection time point for every client in the dataset, as well as nominally missing data because clients varied in the time spent in counselling and therefore the number of time points at which observations were collected. This limited the analyses that I conducted in Study 2, in particular the repeated measures ANOVA, which could only include the data of the sub-sample of clients who had a full set of data across 40 sessions. Missing data was also an issue in Study 3: variations in the wider protocol across time as well as in data collected and retained in individual client files (likely arising from varying degrees of motivation amongst clients, therapists and researchers) resulted in substantial differences in detail available for the rich case records collated by my HSCED researchers, especially for the clients designated as ‘deteriorators’. These

discoveries have reinforced my awareness of the potential for care and attention to lapse during data collection designed for secondary use at an unknown later date, especially when the immediate needs or preferences of vulnerable clients may cause data collection to be de-prioritised.

Generalisability. Following the guidance of the Standards for Educational and Psychological Testing (AERA, APA & NCME, 2014), evidence of validity developed in this thesis should not be generalised for the use of the SI, and the interpretation of its scores, with other client populations or within other contexts and activities in which the development of congruent functioning is promoted. The evidence accrued in this thesis is based on data provided by UK-based clients accessing a free counselling service situated within a university-based research environment. Each of these characteristics may have implications for the potential generalizability of the findings, including assumptions and norms implicit in UK culture and the manifold socio-economic issues that may attract clients to access free counselling, but also deter them from using a service that this is based on a university campus, and require sufficient literacy to take part in research activities.

In addition to this, there are two other specific features of this thesis that may be potential limitations to the generalisability of the findings and therefore should be highlighted: that the therapy was conducted by *trainee therapists in a research environment*, and my decision to study *extreme cases* in the third study.

Trainee therapists in a research environment. Although it is generally understood in counselling research that outcome is not necessarily predicted by therapist experience (Cooper, 2008, p.95), it is certainly possible that trainee therapists learning to practice while working in the conditions required by a research clinic, are affected by the environment. In my own experience – and in listening to the trainees – particularly difficult aspects include: the degree of exposure and scrutiny created by the comprehensive recording of therapy sessions; regular review of the client's experience of the therapy (and therapist) conducted confidentially by a peer of the trainee; and not knowing if and when any particular piece of client work may be analysed by a researcher. The potential impact on

trainee therapists' practice when working in this type of environment is an ethical issue and part of the therapists' process that is supported and explored through supervision as they develop experience and trust in their own abilities. Nevertheless, it is plausible that, at least for some clients and some therapists, the therapy conducted by these trainees may be influenced by a fear of failure that paradoxically inhibited their ability to offer effective person-centred therapy, and it is certainly possible that this effect was demonstrated in the findings of my third study. If this is the case, the extent to which it may impact on the data collected in the full data set is unknown as is the resulting generalisability of the findings.

Studying extreme cases. The meta-synthesis conducted in the third study was based on the in-depth examination of a series of 'extreme' cases: clients whose scores on the Strathclyde Inventory demonstrated reliable change, either indicating improvement or deterioration. However, the majority of clients within the dataset had pre-post scores that suggested non-significant change. Therefore, it is possible that the conclusions of my thesis do not apply to their experience and this requires further investigation.

Piloting adapted methods. In order to fully develop my thesis, I piloted innovations to existing methods: adaptations to the HSCED method, and in the application of procedures associated with grounded theory analysis to conducting a meta-synthesis of a series of case studies. As is typical in pilot studies, implementation of these adaptations required the ability to adjust as ideas that made sense on paper were exposed to experience. In particular, within the HSCED group project, it was necessary for me to make decisions in consultation with the students that enabled them to complete their projects within the time and levels of energy available to them as one strand of an intensive taught postgraduate and practitioner development programme (see Chapter 6). The lessons learned include greater awareness of the time required by adjudicators to digest HSCED cases prior to meeting as an adjudication panel, and the potential influence on adjudicators of preparing an alternatively argued case (i.e. researcher who argued for improvement

then adjudicating a case arguing deterioration). This learning should be taken forward into future projects that plan to conduct and adjudicate multiple HSCEDs.

Researcher allegiance. There is no doubt that the researchers involved in collecting and analysing the data on which this thesis is based had strong allegiance to person-centred therapy. The majority were trainee person-centred therapists. Their views, along with mine, have shaped this thesis at all stages of this process: whether influencing the perceptions of participants during the research process; selecting evidence to include, and arguments to advance, in their HSCED studies; or in choosing research questions, conducting analyses and interpreting results. It is highly likely that another group of researchers, approaching the data from an alternative perspective, may have reached different conclusions. Nevertheless, I have sought to be transparent about the choices that I have made throughout this process so that my reasoning is explicit and contextualised.

Recommendations for Future Research

In this section, I present my recommendations for future studies that take forward specific questions or themes arising from my thesis. These are: (a) ongoing validation of the SI for use in other contexts; (b) further empirical investigation of the construct, congruent functioning, and its developmental pathway; and (c) the use of mixed method approaches in measure development.

Ongoing validation of the Strathclyde Inventory. This thesis has gathered compelling evidence for the validity of using and interpreting scores collected on the Strathclyde Inventory as a measure of outcome in person-centred therapy for UK-based clients accessing counselling in a university-based research environment. Therefore, ongoing work should be conducted in order to test and expand the potential applicability of the instrument with other client populations, contexts, and activities, for example: therapy training, encounter groups, organisational development. Furthermore, it is essential that the validity of the new 12-item version of the SI is tested with newly collected data. This thesis offers a template for future validation studies.

Further investigation of congruent functioning. My results have provided preliminary evidence about the nature of congruent functioning and the process of change that occurs over the course of therapy. Some of the proposed next steps could be carried out using the existing dataset but, for others, new data collection will be required.

First, it is vital to investigate the experience of those clients who did not record clinically significant change, according to their scores on the SI, given the high proportion of clients in the dataset with this outcome. As I outlined in Chapter 5, this is not an unusual finding in counselling research, and therefore it is imperative that we seek to understand what has, or has not, taken place in therapy for these clients. There is an implication that these clients' scores flat-lined, and that therefore no change occurred, however this was found not to be the case for the three clients identified in Chapter 5. Investigating the experience of clients with 'mixed' results is a challenge because it requires the evaluation of contradictory evidence (e.g. Stephen et al., 2011) but one that has a strong likelihood of making a significant contribution to the field. This could be conducted through a replication of my third study: a series of case studies investigating individual outcomes, with common themes identified through a meta-synthesis.

Second, it is necessary to investigate the outcome and process of participants whose pre-therapy scores were in the non-clinical range on the instrument. This has the potential to advance our understanding by exploring the relationship between congruent functioning and distress, and at the same time investigate the possibility that high pre-therapy scores may, in fact, indicate some degree of incongruent functioning at the beginning of therapy. Again, one way to approach this question would be through a replication of my third study, incorporating specific questions about the ways in which distress was experienced and expressed by these clients and how this changed over the course of therapy.

Third, the relationship between a reliable increase or decrease in scores during therapy and the decision to end therapy should be explored further. Does this indicate an awareness for the client that something significant in their

experience has changed? A programme of process-outcome research should be conducted to investigate this phenomenon associated with the *good enough level* model. The existing dataset could provide data for a pilot study in order to develop a method before pursuing new data collection.

Fourth, this thesis has provided a foundation for further investigation of patterns of change in congruent functioning over the course of therapy. Given the apparent points of connection with Pascual-Leone's (2018) model for emotional transformation, and the variety in the shape of change in congruent functioning indicated by my study, it would be very interesting to conduct a study that explores the degree to which Pascual-Leone's *two steps forward, one step back* model (2018, p.170) - a saw-toothed pattern of emotional progress within and across sessions – applies to the development of congruent functioning in therapy. This type of pattern may explain the variation in scores across therapy for many clients, which ultimately results in an insignificant pre-post change in scores.

Finally, this research into the experience of congruent functioning is based on the use of self-report methods. It has provided an alternative perspective on the issue of response shift, previously identified as a limitation in the use of self-report methods. Research is required that explores the apparent link between increasing congruent functioning and response shift in self-report in order to better understand this phenomenon and to develop or revive alternative methods for capturing the experience of congruent functioning that involve multiple perspectives (e.g. client, therapist, observer), such as Elliott et al.'s (1990) development of Walker, Rablen and Roger's (1960) observer-rated *Process Scale* for use as a therapist-rated instrument.

A mixed method approach to measure development. To the best of my knowledge, this is the first measure development study that included the meta-synthesis of a series of systematic case studies. In doing so, I have demonstrated that integrating a mixed method approach of this nature with more conventional measure development processes has the potential to greatly enhance the evidence available to the validation process, providing a solid link between the abstract data

collected by the instrument and the real life experience in therapy of the clients who provided it.

In this thesis, I also introduced developments to the HSCED method itself: in particular, testing an adapted process that better targeted the evaluation of deterioration as outcome, and piloting procedures that supported adjudication panels to make their decision by consensus. The results raised new issues relating to potential adjudicator bias – when adjudicators have recently been immersed in seeking evidence for alternative arguments (i.e. improvement versus deterioration) - that should be considered in the application of these procedures in a future study.

Conducting a meta-synthesis using the products of this series of case studies was an innovation for case study research, as well as for measure development. In practice, I employed an adapted method of grounded theory analysis; future studies should embrace the opportunity to continue the process of refining and developing this method.

Implications for Practice

My findings raise implications for two important aspects of practice: (a) working with doubts, discomfort and disruption; and (b) adopting an informed approach to measurement in counselling.

Working with Doubts, Discomfort and Disruption

This thesis has provided evidence that doubts, discomfort and disruption, experienced by both clients and therapists, restrict the development of congruent functioning within the therapeutic process. This finding has implications for practice, supervision and training.

First, counsellors need to be aware that clients are likely to hold doubts (or expectations that could lead to doubt) about the therapeutic process. If unattended, it is probable that these doubts will be reinforced by the inevitable discomfort experienced by these clients in therapy. In turn, this discomfort will lead to discomfort within the therapist (e.g doubt and frustration) that impacts on their capacity to work therapeutically. This is a vicious circle within the therapeutic

relationship that is unlikely to lead to a beneficial outcome for the client. Therefore, therapists need to be sufficiently developed (i.e. in their capacity for empathic attunement and congruent functioning) to be able to recognise what is occurring between themselves and the client, and prepared to open up and hold 'difficult' conversations in which they seek to understand and accept the client along with their doubts and discomfort. Therapists need to be aware of the likelihood that clients may respond with deference (Rennie, 1994) in the moment because they may not yet trust that their experience will be understood or may not yet be in touch with its impact or meaning. Furthermore, therapists need to be aware of their own reactions to the client: what triggers are being pressed in them?, and what impact is this having on their response to the client? In order to support therapists, supervisors should listen out for these potential difficulties, perhaps unspoken, within the therapeutic process, and trainers should ensure that sufficient attention is paid within the curriculum to this challenging aspect of practice.

Second, a change of therapist mid-process will have an unavoidable impact on the client's development. However, it is not inevitable that this will lead to a negative outcome for the client if sufficient attention is paid to the potential impact of the change on the client (in particular the timing) and to building a new therapeutic relationship that can be available until the end of the client's therapeutic process. This is vital: my findings suggest that a further change in therapist when the client is still mid-process is highly likely to lead to loss of hope for the client and the decision (formally or informally) to leave therapy. This raises serious considerations for counsellors and counselling services. Despite best intentions, it is not always possible to guarantee that a counsellor will be available to work to the end of their client's process, especially in services that offer longer-term or open-ended counselling. There is no easy solution to this issue but the first step is to highlight it as a major challenge for clients, therapists and service providers, which requires a well-considered response.

Adopting an Informed Approach to Measurement in Counselling

As I outlined in Chapter 3, the practice of measurement in counselling has the potential to be misused and even abused when its limitations are misunderstood. As my findings indicate, representing change by simply calculating the difference between pre- and post-therapy scores is not as straight-forward as often believed. This is for several reasons, for example: a non-clinical pre-therapy score, response shift, shape of change across therapy, and (for group data) the phenomenon of effect size overestimating the experience of individual clients when measured as clinically significant change (Lambert, 2013, p.178). Nevertheless, counsellors and counselling services alike tend to assume that this is a sufficient and meaningful way to extract information about outcome from their data. I have worked with supervisees who feel anxious and embarrassed about the potential implications of their clients' outcome data.

My thesis demonstrates the value of evaluating measurement data within its wider context in which there is an attempt to understand what has occurred between the beginning and end of therapy. While this is not possible for all client data in all contexts, it provides a framework for making the best use of measurement *for* counselling. The data does not necessarily predict or explain clients' experiences and outcomes, but provides a snapshot of current progress (or capacity for congruent functioning, in the case of the Strathclyde Inventory) and a means to identify cases (e.g. clients whose scores indicate extreme change, whether improvement or deterioration, or whose scores fluctuate or flat-line) that therapists can use to focus and explore their learning and development within supervision (i.e. *evidence-based reflective practice*; Lucock et al., 2003), and counselling services can use to expand their understanding of the challenges and outcomes experienced by clients and counsellors working together within their organisation.

Conclusion

This comprehensive investigation has demonstrated that congruent functioning, and change in congruent functioning as an outcome of person-centred therapy, can be measured using the Strathclyde Inventory. It has: (a) confirmed that

person-centred theory, originally developed through cutting edge research in the mid-twentieth century, continues to be supported using data drawn from contemporary practice; (b) extended this theory by using a mixed methods approach to validation that has identified a developmental pathway for congruent functioning as well as interfering therapeutic processes that inhibit growth; and (c) highlighted a complementary relationship between the core theory underpinning person-centred therapy, emotion-focused therapy and contemporary research in self-compassion and emotional transformation. These findings have major implications for theory, research and practice, which I have presented in this chapter; now I conclude with a summary of the original contributions to knowledge made by this thesis and share some reflections on my learning.

Original Contributions to Knowledge

- An evidence base supporting the validity of using the Strathclyde Inventory as a therapy outcome measure with a UK-based clinical population.
- A brief 12-item version of the Strathclyde Inventory.
- Clear empirical evidence that supports and extends the nomological net underpinning congruent functioning.
- In particular, an evidence-based developmental pathway for congruent functioning, with self-acceptance as a detectable pivot point.
- Key features of the therapeutic process that facilitate and inhibit outcome, with implications for therapists, trainers, supervisors, and service providers.
- An innovative mixed method approach to measure development that grounds validation in practice.
- An evidence base for adopting an informed approach to measurement in counselling.

Reflections on my Learning

Finally, what have I learned during this process, as a person, a practitioner and a researcher?

First, completing this thesis provided me with the opportunity to resonate with the developmental pathway for congruent functioning in my own process as I encountered the challenges involved. I experienced self-doubt and painful awareness of my own limitations. I was fortunate to be supported by supervisors, colleagues and friends who listened and understood, enabling me to stay open to my experience. I felt a turning point when I became able to let go of my preconceptions of who and what I thought I needed to be to complete a PhD thesis, and to value the qualities and interests that I have. Receiving positive feedback from people that I trusted, enabled me to take the risk of trusting myself. I began to recognise that my thesis is a reflection of the person that I am, not anyone else, and that this is good enough. Through this experience, I found my flow in the process, becoming more and more ready and willing to share the evolution of my thesis with others.

How has this influenced my practice as a therapist, supervisor, trainer and service coordinator? Already part of my practice, it has reinforced my commitment to open dialogue (i.e. meta-communication) with my clients about our work. I am increasingly responsive to ambivalence that I pick up from clients, and between us, especially as we begin, having learned from this research that doubts are likely to play out in the process with detrimental consequences. I encourage my supervisees to do the same. Indeed, it is becoming typical in my supervision practice for me to ask my supervisees: have you checked this out with your client? We have not yet reached a point where they anticipate my question! This raises concerns about the attention paid to these issues in training. I believe that the learning from my thesis, especially the features of therapy that can lead to potential deterioration, has much to contribute to person-centred training curricula. Whether or not I return to a substantive role in training, it is my intention to encourage peers and colleagues to focus explicitly on this area within their own courses. Finally, as the former coordinator of a counselling service, I recognise the challenge of achieving consistency for clients especially when offering a longer-term service. In some ways self-evident, these findings are nevertheless important as they demonstrate the

potential impact of a change in therapist on client outcome, which can be minimised or overlooked in the context of a busy service. I hope to share this learning and to work with services to develop ways to respond to this challenge.

An unexpected outcome of this thesis is my interest in creating opportunities to continue working with measurement. This is surprising to me as before beginning this work I viewed measurement as a means to an end. On reflection, this is not surprising: as a counsellor I was immersed in an experience-near perspective. Indeed, for some time during this work I found myself adopting an apologetic stance when describing my research to colleagues. It was interesting to discover that when I described my findings, the majority of colleagues became very interested indeed! Through this experience, my stance has shifted: while I recognise its limitations as a method, I have come to appreciate the considerable value and potential contribution of measurement when used in an informed way. As a result, I want to use my learning to make an innovative and timely contribution within the field by actively working to make measurement in counselling more useful and user-friendly, and to enable counsellors to develop their own informed approach to measurement that integrates an understanding of its limitations and also its potential.

References

Alhimaidi, A. (2019). *The Nature and Effectiveness of Therapies Provided in Counselling Centres at the Princess Nourah bint Abdulrahman University.* Unpublished PhD thesis, University of Strathclyde.

American Educational Research Association, American Psychological Association, & National Council on Measurement in Education. (2014). *Standards for educational and psychological testing*. Washington, DC: American Educational Research Association.

Amrhein, V., Greenland, S., & McShane, B. (2019). *Scientists rise up against statistical significance*. Retrieved from: https://www.nature.com/articles/d41586-019-00857-9

Ashby, J.S., Rahotep, S.S., & Martin, J.L. (2005). Multidimensional perfectionism and Rogerian personality constructs. *Journal of Humanistic Counseling, Education and Development, 44,* 55-65.

BACP (2019). *Ethical Guidelines for Research in the Counselling Professions.* Lutterworth: BACP.

Bagby, R.M., Parker, J.D.A., & Taylor, G.J. (1994). The twenty-item Toronto Alexithymia Scale – I. Item selection and cross-validation of the factor structure. *Journal of Psychosomatic Research, 38*, 23-32.

Barkham, M., Connell, J., Stiles, W.B., Miles, J.N.V., Margison, F., Evans, C., & Mellor-Clark, J. (2006). Dose-Effect Relations and Responsive Regulation of Treatment Duration: The Good Enough Level. *Journal of Consulting & Clinical Psychology, 74,* 160-167.

Barkham, M., Margison, F., Leach, C., Lucock, M., Mellor-Clark, J., Evans, C., Benson, L., Connell, J., Audin, K., & McGrath, G. (2001). Service Profiling and Outcomes Benchmarking Using the CORE-OM: Toward Practice-Based Evidence in the Psychological Therapies. *Journal of Consulting and Clinical Psychology, 69*, 184-196.

Barkham, M. & Mellor-Clark, J. (2003). Bridging Evidence-Based Practice and Practice-Based Evidence: Developing a Rigorous and Relevant Knowledge for the Psychological Therapies. *Clinical Psychology & Psychotherapy, 10,* 319-327.

Barker, C., Pistrang, N. & Elliott, R. (2016). *Research Methods in Clinical Psychology*. Chichester: Wiley.

Barnett, M.D., Moore, J.M., & Harp, A.R. (2017). Who we are and how we feel: Self-discrepancy theory and specific affective states. *Personality and Individual Differences, 111,* 232-237.

Barnett, M.D. & Womack, P.M. (2015). Fearing, not loving, the reflection: Narcissism, self-esteem, and self-discrepancy theory. *Personality and Individual Differences, 74,* 280-284.

Barrett-Lennard, G.T. (1998). *Carl Rogers' Helping System: Journey and Substance.* London: Sage.

Bauer, S., Lambert, M.J., & Nielsen, S.L. (2004). Clinical Significance Methods: A Comparison of Statistical Techniques. *Journal of Personality Assessment, 82,* 60-70.

Behr, M. & Becker, M. (2002). Congruence and experiencing emotions: Self-report scales for the person-centered and experiential theory of personality. In J. C. Watson, R. N. Goldman, M. S. Warner (Eds.), *Client-centered and experiential psychotherapy: Advances in theory, research and practice* (pp. 150-167). Ross-on-Wye: PCCS Books.

Behr, M. & Becker, M. (2012). Scales for Experiencing Emotions: Awareness, appraisal and regulation of one's own emotions. *Hellenic Journal of Psychology, 9,* 278-303.

Benelli, E., De Carlo, A., Biffi, D. & McLeod, J. (2015). Hermeneutic Single Case Efficacy Design: A systematic review of published research and current

standards. *Testing, Psychometrics, Methodology in Applied Psychology, 22,* 97-133.

Bishop, S.L., Walling, D.P., & Walker, B.A. (1997). The Emperor's Clothes: Assessing the Validity of Scores on the Tennessee Self-Concept Scale. *Educational and Psychological Measurement, 57,* 150-163.

Blanton, H. & Jaccard, J. (2006). Arbitrary Metrics in Psychology. *American Psychologist, 61*, 27-41.

Blount, C., Evans, C., Birch, S., Warren, F. & Norton, K. (2002). The properties of self-report research measures: Beyond psychometrics. *Psychology and Psychotherapy: Theory, Research and Practice, 75,* 151-164.

Bond, M.J., Strauss, N.E., & Wickham, R.E. (2018). Development and validation of the Kernis-Goldman authenticity inventory short form (KGAI-SF). *Personality and Individual Differences, 134,* 1-6.

Bond, T.G. & Fox, C.M. (2015). *Applying the Rasch Model: Fundamental Measurement in the Human Sciences*. Hove: Routledge.

Bondi, L. & Fewell, J. (2016). *Practitioner Research in Counselling and Psychotherapy.* London: Palgrave Macmillan.

Boswell, J.F., Kraus, D.R., Miller, S.D., & Lambert, M.J. (2015). Implementing routine outcome monitoring in clinical practice: Benefits, challenges, and solutions. *Psychotherapy Research, 25,* 6-19. DOI: 10.1080/10503307.2013.817696

Bozarth, J.D. (2001a). Congruence: A special way of being. In G. Wyatt (Ed.) *Congruence* (pp.184-199). Ross-on-Wye: PCCS Books.

Bozarth, J.D. (2001b). Client-centered Unconditional Positive Regard: A historical perspective. In J.D. Bozarth & P. Wilkins (Eds.) *Unconditional Positive Regard* (pp.5-18). Ross-on-Wye: PCCS Books.

Boyraz, G., & Kuhl, M.L. (2015). Self-focused attention, authenticity, and well-being. *Personality and Individual Differences, 87,* 70-75.

Brainscape (n.d.) *Lecture 2 Theoretical perspectives on aging well.* Retrieved from: https://www.brainscape.com/flashcards/lecture-2-theoretical-perspectives-on-agi-5373526/packs/8085084.

Brattland, H., Koksvik, J.M., Burkeland, O., Gråwe, R.W., Klöckner, C., Linaker, O.M., Ryam, T., Wampold, B., Lara-Cabrera, M.L., & Iversen, V.C. (2018). The Effects of Routine Outcome Monitoring (ROM) on Therapy Outcomes in the Course of an Implementation Process: A Randomized Clinical Trial. *Journal of Counseling Psychology, 65,* 641-652. DOI:10.1037/cou0000286

Butler, J.M. & Haigh, G.V. (1954). Changes in the relation between self-concepts and ideal concepts consequent upon client-centered counseling. In C.R. Rogers & R.F. Dymond (Eds.) *Psychotherapy and Personality Change: Coordinated research studies in the client-centered approach* (pp.55-75). Chicago, IL: The University of Chicago Press.

Carrick, L., & Elliott, R. (2013). *The therapeutic relationship in Person-Centred-Experiential Psychotherapy: Measurement and change over time*. Paper presented at conference of the UK Chapter of the Society for Psychotherapy Research, Oxford, UK.

Cartwright, D., deBruin, J. & Berg, S. (1991). Some Scales for Assessing Personality Based on Carl Rogers' Theory: Further Evidence of Validity. *Personality and Individual Differences, 12,* 151-156.

Cartwright, D. & Mori, C. (1988). Scales for assessing aspects of the person. *Person-Centered Review, 3*, 176-194.

Cartwright, D. & Peckar, H. (1993). Purposefulness: A fourth superfactor? *Personality and Individual Differences, 14*, 547-555.

Cohen, J. (1992). A Power Primer. *Psychological Bulletin, 112*, 155-159.

Comer, J.S. & Kendall, P.C. (2013). Methodology, Design and Evaluation in Psychotherapy Research. In M. Lambert (Ed.) *Bergin & Garfield's Handbook of Psychotherapy and Behavior Change* (pp.21-48). New York: Wiley.

Connell, J., Barkham, M., Stiles, W.B., Twigg, E., Singleton, N., Evans, O., & Miles, J.N.V. (2007). Distribution of CORE-OM scores in a general population, clinical cut-off points and comparison with the CIS-R. *British Journal of Psychiatry, 190,* 69-74.

Cook, T.D. & Campbell, D.T. (1979). *Quasi-Experimentation: Design & Analysis Issues for Field Settings.* Boston: Houghton Mifflin.

Cooper, M. (2008). *Essential Research Findings in Counselling and Psychotherapy.* London: Sage.

Cooper, M., O'Hara, M., Schmid, P.F., & Bohart, A.C. (Eds.) (2013) *The Handbook of Person-Centred Psychotherapy & Counselling.* Basingstoke: Palgrave Macmillan.

Corbin, J. & Strauss, A. (2008). *Basics of Qualitative Research (3rd Ed.).* Thousand Oaks, CA: Sage.

CORE System Trust (n.d.). *CORE Outcome Measure*. Rugby: CORE IMS Ltd.

Cornelius-White, J.H.D. (2007). Congruence: An integrative five-dimensional model. Person-Centred & Experiential Psychotherapies, 6, 229-239.

Cornette, M.M., Strauman, T.J., Abramson, L.Y., & Busch, A.M. (2009). Self-discrepancy and suicidal ideation. *Cognition and Emotion, 23,* 504-527.

Costa, P.T., Jr., and McCrae, R.R. (1992). *Revised NEO Personality Inventory (NEO PI-R) and the NEO Five-Factor Inventory (NEO-FFI) professional manual*. Odessa, FL: Psychological Assessment Resources.

Crandal, R. (1973). The measurement of self-esteem and related constructs. In J. P. Robinson & P. R. Shaver (Eds.), *Measures of social psychological attitudes* (pp.80-82). Ann Arbor: ISR.

Cronbach, L.J. & Furby, L. (1970). How we should measure “change” – or should we? *Psychological Bulletin, 74,* 68-80.

Cronbach, L.J. & Meehl, P.E. (1955). Construct validity in psychological tests. *Psychological Bulletin, 52,* 281-302.

Curran, J., Parry, G.D., Hardy, G.E., Darling, J., Mason, A-M., & Chambers, E. (2019). How Does Therapy Harm? A Model of Adverse Process Using Task Analysis in the Meta-Synthesis of Service Users’ Experience. *Frontiers in Psychology, 10*, art.347. DOI: 10.3389/fpsyg.2019.00347

Deci, E.L. & Ryan, R.M. (2000). The "What" and "Why" of Goal Pursuits: Human Needs and the Self-Determination of Behavior. *Psychological Inquiry, 11,* 227-268.

Delgadillo, J., Asaria, M., Ali, S., & Gilbody, S. (2016). On poverty, politics and psychology: the socioeconomic gradient of mental healthcare utilisation and outcomes. *The British Journal of Psychiatry, 209,* 429-430.

De Vellis, R.F. (2017). *Scale Development: Theory and Applications.* Thousand Oaks, CA: Sage.

Eckert, J., Höger, D., & Schwab, R. (2003). Development and Current State of the Research on Client-Centered Therapy in the German Language Region. *Person-Centered & Experiential Psychotherapies, 2,* 3-18.

Eisen, S.V., Ranganathan, G., Seal, P., & Spiro, A. (2007). Measuring Clinically Meaningful Change Following Mental Health Treatment. *Journal of Behavioral Health Services & Health, 34,* 272-289.

Ellingham, I. (2001). Congruence as an organismic, not a Freudian concept. In G. Wyatt (Ed.) *Congruence*. Ross-on-Wye: PCCS Books.

Elliot, N. (2015). Validation: The pursuit. *College Composition and Communication; Urbana, 66,* 668-687.

Elliott, R. (2001). The effectiveness of humanistic therapies: A meta-analysis. In D. Cain & J. Seeman (Eds.) *Humanistic psychotherapies: Handbook of research and practice* (pp. 57-82). Washington, D.C.: APA.

Elliott, R. (2002). Render unto Caesar: Quantitative and qualitative knowing in research on humanistic therapies. *Person-Centered & Experiential Psychotherapies, 1,* 102-117.

Elliott, R. (2004). *How to Measure Clinically Significant Change*. Unpublished training material, University of Toledo, OH, USA.

Elliott, R. (2009). *Twenty-Five Fun Research Projects to Do in a Therapy Research Clinic*. Retrieved from https://pe-eft.blogspot.com/2009/09/twenty-five-fun-research-projects-to-do.html

Elliott, R. (2014). Hermeneutic Single-Case Efficacy Design: An Overview. In K.J.Schneider, J.F. Fraser & J.F.T. Bugental (Eds.), *Handbook of Humanistic Psychology: Leading Edges in Theory, Practice & Research* (2nd Ed.) (pp. 351-360). Thousand Oaks, CA: Sage.

Elliott, R. (2015). Hermeneutic single case efficacy design (2002). Looking back: commentary on “Hermeneutic single case efficacy design”. In Strauss, B., Barber, J.P., & Castonguay, L.G. (Eds.), *Visions in Psychotherapy Research and Practice: Reflections from the Presidents of the Society for Psychotherapy Research* (pp. 188-208). London: Routledge.

Elliott, R., Clark, C., Kemeny, V., Wexler, M.M., Mack, C., & Brinkerhoff, J. (1990). The Impact of Experiential Therapy on Depression: Initial Results. In G. Lietaer, J. Rombauts, & R. Van Balen (Eds.), *Client-centered and experiential psychotherapy towards the nineties* (pp.549-577). Leuven, Belgium: Leuven University Press.

Elliott, R., Fox, C.M., Beltyukova, S.A., Stone, G.E., Gunderson, J. & Zhang, X. (2006). Deconstructing therapy outcome measurement with Rasch analysis: The SCL-90-R. *Psychological Assessment, 18*, 359-372.

Elliott, R., Greenberg, L.S., Watson, J., Timulak, L. & Freire, E.S. (2013). Research on Humanistic-Experiential Psychotherapies. In M.J. Lambert (Ed.) *Bergin & Garfield's Handbook of Psychotherapy and Behavior Change* (pp.495-538). Hoboken, N.J.: John Wiley & Sons, Inc.

Elliott, R., Partyka, R., Alperin, R., Dobrenski, R., Wagner, J., Messer, S., Watson, J.C., & Castonguay, L.J. (2009). An adjudicated hermeneutic single-case efficacy design study of experiential therapy for panic/phobia. *Psychotherapy Research, 19,* 543-557.

Elliott, R. & Rodgers, B. (2007). *Strathclyde Inventory - 31 item version (Version 4).* Unpublished instrument, University of Strathclyde, Glasgow, UK.

Elliott, R. & Rodgers, B. (2012a). *Strathclyde Inventory – 31 item version (Version 4).* Unpublished data analyses, University of Strathclyde, Glasgow, UK.

Elliott, R. & Rodgers, B. (2012b). *Strathclyde Inventory – 16 item version (Version 6).* Unpublished data analyses, University of Strathclyde, Glasgow, UK.

Elliott, R. & Rodgers, B. (2012c). *Strathclyde Inventory - 16 item version (Version 6).* Unpublished instrument, University of Strathclyde, Glasgow, UK.

Elliott, R., Rodgers, B., Stephen, S. (2013). *Person-Centred vs Emotion-Focused Therapies for Social Anxiety: Outcome Results*. Paper presented at the meeting of the Society for Psychotherapy Research, Brisbane, Australia.

Elliott, R., Slatick, E., & Urman, M. (2001). Qualitative change process research on psychotherapy: Alternative strategies. In J. Frommer & D.L. Rennie (Eds.), *Qualitative Psychotherapy Research: Methods and Methodology*. Lengerich, Germany: Pabst Science Publishers.

Elliott. R., Wagner, J., Sales, C.M.D, Rodgers, B., Alves, P., & Café, M.J. (2016). Psychometrics of the Personal Questionnaire: A Client-Generated Outcome Measure. *Psychological Assessment, 28,* 263-278.

Elliott, R., Watson, J.C., Goldman, R.N. & Greenberg, L.S. (2004). *Learning Emotion-Focused Therapy*. Washington D.C.: American Psychological Association.

Evans, C. (2019). *Routine Outcome Monitoring (ROM): A great tool but people aren't silicon!* Online webinar hosted on 27 March 2019 by the Society for Psychotherapy Research.

Evans, C., Connell, J., Barkham, M., Margison, F., McGrath, G., Mellor-Clark, J. & Audin, K. (2002). Towards a standardised brief outcome measure: Psychometric properties and utility of the CORE-OM. *British Journal of Psychiatry, 180*, 51-60.

Evans, C., Connell, J., Barkham, M., Marshall, C., & Mellor-Clark, J. (2003). Practice-Based Evidence: Benchmarking NHS Primary Care Counselling Services at National and Local Levels. *Clinical Psychology & Psychotherapy, 10,* 374-388.

Faraci, P. & Cannistraci, S. (2015). The Short Index of Self-Actualization: A factor analysis study in an Italian sample. *International Journal of Psychological Research, 8,* 23-33.

Farr, J., Di Malta, G., & Cooper, M. (2019). Pilot randomised controlled trial of counselling in a community mental health setting: Pitfalls and learnings. *Counselling & Psychotherapy Review.* DOI: 10.1002/capr.12262

Felton, K. (2005). Meaning-based quality-of-life measurement: A way forward in conceptualizing and measuring client outcomes? *British Journal of Social Work, 35,* 221-236.

Field, A. (2013). *Discovering Statistics Using IBM SPSS Statistics.* London: Sage.

Flanagan, S., Patterson, T.G., Hume, I.R., & Joseph, S. (2015). A longitudinal investigation of the relationship between unconditional positive self-regard and posttraumatic growth. *Person-Centered & Experiential Psychotherapies, 14,* 191-200.

Fokkema, M. & Greiff, S. (2017). How Performing PCA and CFA on the Same Data Equals Trouble: Overfitting in the Assessment of Internal Structure and Some Editorial Thoughts on It. *European Journal of Psychological Assessment, 33*, 399-402. DOI: 10.1027/1015-5759/a000460

Folkes-Skinner, J.A. (2011). *A Mixed Method Study of How Trainee Counsellors Change.* Unpublished PhD Dissertation, University of Leicester.

Freire, E. S. (2007). *The Strathclyde Inventory: A psychotherapy outcome measure based on the person-centred theory of change*. Unpublished MSc dissertation, University of Strathclyde.

Freire, E S. (2009). *Strathclyde Inventory – 22 item version.* Unpublished data analyses, University of Strathclyde, Glasgow, UK.

Freire, E.S., Elliott, R., & Westwell, G. (2014). Person-Centred and Experiential Psychotherapy Scale: Development and reliability of an adherence/ competence measure for person-centred and experiential psychotherapies. *Counselling & Psychotherapy Research, 14,* 220-226.

Freire, E.S., Williams, C., Messow, C-M, Cooper, M., Elliott, R., McConnachie, A., Walker, A., Heard, D., & Morrison, J. (2015). Counselling versus low-intensity cognitive behavioural therapy for persistent sub-threshold and mild depression (CLICD): a pilot/feasibility randomised controlled trial. *BMC Psychiatry, 15,* 197. DOI 10.1186/s12888-015-0582-y

Frick, W.B. (1999). Flight into Health: A New Interpretation. *Journal of Humanistic Psychology, 39,* 58-81.

Galili-Weinstock, L., Chen, R., Atzil-Slonim, D., Bar-Kalifa, E., Peri, T., & Rafaeli, E. (2018). The association between self-compassion and treatment outcomes: Session-level and treatment-level effects. *Journal of Clinical Psychology, 74,* 849-866.

Gendlin, E.T. (1996). *Focusing-oriented Psychotherapy.* New York, NY: The Guilford Press.

Giddens, A. (1984). *The constitution of society: Outline of the theory of structuration.* Cambridge: Polity Press.

Gold, B.T. (1996). Enviousness and its relationship to maladjustment and psychopathology. *Personality and Individual Differences, 21*, 311-321.

Goldman, B.M. & Kernis, M.H. (2002). The role of authenticity in healthy psychological functioning and subjective well-being. *Annals of the American Psychotherapy Association, 5(6),* 18-20.

Goldman, R.N., Greenberg, L.S., & Pos, A.E. (2005). Depth of emotional experience and outcome. *Psychotherapy Research, 15,* 248-260.

Gorsuch, R.L. (1997). Exploratory factor analysis: Its role in item analysis. *Journal of Personality Assessment, 68*, 532-560.

Grafanaki, S. (2013). Experiencing congruence and incongruence. *Person-Centered & Experiential Psychotherapies, 12,* 183-186.

Grant, B. (1990). Principled and instrumental nondirectiveness in person-centered and client-centered therapy. *Person-Centered Review, 5,* 77-87.

Grégoire, S., Baron, L., Ménard, J. & Lachance, L. (2014). The Authenticity Scale: Psychometric Properties of a French Translation and Exploration of Its Relationships With Personality and Well-Being. *Canadian Journal of Behavioural Science, 46,* 346-355.

Griffiths, L.J., & Griffiths, C.A. (2013). Unconditional Positive Self-Regard (UPSR) and Self-Compassion, the Internal Consistency and Convergent/Divergent Validity of Patterson & Joseph's UPSR Scale. *Open Journal of Medical Psychology, 2,* 168-174.

Grijak, D. (2017). Psychometric evaluation of the Authenticity Scale on the sample of students in Serbia. *Psihologija, 50,* 85-99.

Gyani, A., Shafran, R., Myles, P., & Rose, S. (2014). The Gap Between Science and Practice: How Therapists Make Their Clinical Decisions. *Behavior Therapy, 45*, 199-211.

Hardin, E.E. & Lakin, J.L. (2009). The Integrated Self-Discrepancy Index: A Reliable and Valid Measure of Self-Discrepancies. *Journal of Personality Assessment, 91,* 245-253.

Hardy, G.E., Bishop-Edwards, L., Chambers, E., Connell, J., Dent-Brown, K., Kothari, G., O'Hara, R., & Parry, G.D. (2019). Risk factors for negative experiences during psychotherapy. *Psychotherapy Research, 29,* 403-414. DOI: 10.1080/10503307.2017.1393575

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

Hatfield, D.R., & Ogles, B.M. (2007). Why Some Clinicians Use Outcome Measures and Others Do Not. *Administration and Policy in Mental Health and Mental Health Services Research, 34,* 283-291.

Haugh, S. (2001a). A Historical Review of the Development of the Concept of Congruence in Person-Centred Theory. In G. Wyatt (Ed.) *Congruence* (pp.1-17). Ross-on-Wye: PCCS Books.

Haugh, S. (2001b). The difficulties in the conceptualisation of congruence: A way forward with complexity theory? In G. Wyatt (Ed.) *Congruence* (pp.116-130). Ross-on-Wye: PCCS Books.

Higgins, T.E. (1987). Self-Discrepancy: A Theory Relating Self and Affect. *Psychological Review, 94,* 319-340.

Higgins, T.E., Bond, R.N., Klein, R. & Strauman, T. (1986). Self-Discrepancies and Emotional Vulnerability: How Magnitude, Accessibility, and Type of Discrepancy Influence Affect. *Journal of Personality and Social Psychology, 51,* 5-15.

Higgins, E.T., Klein, R., & Strauman, T. (1985). Self-concept theory: A psychological model for distinguishing among different aspects of depression and anxiety. *Social Cognition, 3,* 51-76.

Hill, C.E., Knox, S., Thompson, B.J., Williams, E.N., Hess, S.A., Ladany, N. (2005). Consensual Qualitative Research: An Update. *Journal of Counseling Psychology, 52,* 196-205.

Hilton, J. & Prior, S. (2018). Re-visioning person-centred research. In M. Bazzano (Ed.) *Re-visioning Person-Centred Therapy. Theory and practice of a radical paradigm* (pp.277-288). London: Routledge.

Holden, R.R. & Marjanovic, Z. (2011). Reaching for the Brass Ring of Psychometric Test Standards: Commenting on Slaney, Storey, and Barnes. *International Journal of Forensic Mental Health, 10,* 284-289.

Howard, K.I., Kopta, S.M., Krause, M.S., & Orlinksky, D.E. (1986). The Dose-Effect Relationship in Psychotherapy. *American Psychologist, 41,* 159-164.

Iwakabe, S. & Gazzola, N. (2009). From single-case studies to practice-based knowledge: Aggregating and synthesizing case studies. *Psychotherapy Research, 19,* 601-611.

Jackson, C. (May 2019). Slaves to the Machine. *Therapy Today, 30*, 8-12. https://www.bacp.co.uk/bacp-journals/therapy-today/2019/may-2019/slaves-to-the-machine/. Accessed 14 June 2019.

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

Jones, C. (2010). Archival Data: Advantages and Disadvantages for Research in Psychology. *Social and Personality Psychology Compass, 4,* 1008-1017. DOI: 10.1111/j.1751-9004.2010.00317.x

Jones, A. & Crandall, R. (1986). *Personality & Social Psychology Bulletin, 12,* 63-73.

Joseph, S. (2016). *Authentic: How to be yourself and why it matters.* London: Piatkus.

Joseph, S. (Ed.) (2017). *The Handbook of Person-Centred Therapy and Mental Health: Theory, Research and Practice.* Monmouth: PCCS Books Ltd.

Kazdin, A.E. (2006). Arbitrary Metrics: Implications for Identifying Evidence-Based Treatments. *American Psychologist, 61,* 42-49.

Kelley, H.H. (1973). The processes of causal attribution. *American Psychologist, 28,* 107-128.

Kennedy-Martin, T., Curtis, S., Faries, D. & Robinson, S., & Johnston, J. (2015). A literature review on the representativeness of randomized controlled trial samples and implications for the external validity of trial results. *Trials, 16:495.* DOI 10.1186/s13063-015-1023-4

Keys, S. (2003). *Idiosyncratic Person-Centred Therapy.* Ross-on-Wye: PCCS Books Ltd.

Klein, M.H., Mathieu-Coughlan, P., & Kiesler, D.J. (1986). The Experiencing Scales. In L.S. Greenberg & W.M. Pinsof (Eds.) *The Psychotherapeutic Process: A Research Handbook* (pp.21-72). New York: The Guilford Press.

Knapp, R.R. & Knapp, L. (1978). Conceptual and statistical refinement and extension of the measurement of actualizing concurrent validation of the Personal Orientation Dimensions (POD). *Educational and Psychological Measurement, 38,* 523-526.

Knoll, M., Meyer, B., Kroemer, N.B., & Schröder-Abé, M. (2015). It Takes Two to Be Yourself. An Integrated Model of Authenticity, its Measurement, and its Relationship to Work-Related Variables. *Journal of Individual Differences, 36,* 38-53.

Kuba, Y. (2013). Congruence and incongruence as human attitudes. *Person-Centered & Experiential Psychotherapies, 12*, 200-208.

Lambert, M.J. (2013). The efficacy and effectiveness of psychotherapy. In M.J. Lambert (Ed.). *Bergin & Garfield's Handbook of Psychotherapy & Behavior Change* (pp.169-218). Hoboken, N.J.: John Wiley & Sons, Inc.

Lambert, M.J., Fidalgo, L.G., & Greaves, M.R. (2016). In D. J. Cain, K. Keenan, and S. Rubin (Eds*.) Humanistic Psychotherapies: Handbook of Research and Practice* (pp.49-79). Washington D.C.: American Psychological Association.

Lambert, M.J., Whipple, J.L., Vermeersch, D.A., Smart, D.W., Hawkins, E.J., Nielsen, S.L., & Goates, M. (2002). Enhancing Psychotherapy Outcomes via Providing Feedback on Client Progress: A Replication. *Clinical Psychology & Psychotherapy, 9,* 91-103.

Langkaas, T.F., Wampold, B.E., & Hoffart, A. (2018). Five Types of Clinical Difference to Monitor in Practice. *Psychotherapy, 55,* 241-254. DOI:10.1037/pst0000194

Lee, D. (2018). Enter centre stage, the case study… In M. Bazzano (Ed.) *Re-visioning Person-Centred Therapy. Theory and practice of a radical paradigm* (pp.211-225). London: Routledge.

Levitt, H.M., Lu, E.C., Pomerville, A. & Francisco, I.S. (2015). Pursuing the question of reflexivity in psychotherapy and qualitative methods: The contributions of David L. Rennie. *Counselling & Psychotherapy Research, 15,* 3-11. DOI: 10.1002/capr.12021

Levitt, H., Stanley, C.M., Frankel, Z., & Raina, K. (2005). An evaluation of outcome measures used in humanistic psychotherapy research: Using thermometers to weigh oranges. *The Humanistic Psychologist*, 33, 113-130.

Linacre, J.M. (1999). Investigating rating scale category utility. *Journal of Outcome Measurement, 3,* 103-122.

Linacre, J. M. (2006). *WINSTEPS Rasch measurement computer program.* Chicago: Winsteps.com

Linacre, J.M. (2010). When to stop removing items and persons in Rasch misfit analysis? *Rasch Measurement Transactions*, 23, 1241. Retrieved from https://www.rasch.org/rmt/rmt234g.htm.

Linacre, J. M. (2011). *Practical Rasch Measurement – Further Topics. Tutorial 3.* Retrieved from http://www.winsteps.com/a/winsteps-tutorial-further-3.pdf.

Linacre, J. M. (2012). *Winsteps Tutorial 4*. Retrieved from http://www.winsteps.com/a/winsteps-tutorial-4.pdf.

Linacre, J. M. (2017). A User's Guide to *WINSTEPS Ministep Rasch-Model Computer Programs: Program Manual 4.0.0).* Chicago: Winsteps.com

Linacre, J.M. & Wright, B.D. (1994). Reasonable mean-square fit values. *Rasch Measurement Transactions, 8,* 370

Llewelyn, S. (1988). Psychological Therapy as viewed by Clients and Therapists. *British Journal of Clinical Psychology, 27*, 223-238.

Lucock, M., Leach, C., Iveson, S., Lynch, K., Horsefield, C., & Hall, P. (2003). A Systematic Approach to Practice-Based Evidence in a Psychological Therapies Service. *Clinical Psychology & Psychotherapy, 10,* 389-399.

Lutz, W., Stulz, N. & Köck, K. (2009). Patterns of early change and their relationship to outcome and follow-up among patients with major depressive disorders. *Journal of Affective Disorders, 118,* 60-68.

Lutz, W., Hofmann, S.G., Rubel, J., Boswell, J.F., Shear, M.K., Gorman, J.M., Woods, S.W., & Barlow, D.H. (2014). Patterns of Early Change and Their Relationship to Outcome and Early Treatment Termination in Patients With Panic Disorder. *Journal of Consulting & Clinical Psychology, 82,* 287-297.

Lyne, K.J., Barrett, P., Evans, C. & Barkham, M. (2006). Dimensions of variation on the CORE-OM. *British Journal of Clinical Psychology, 45,* 185-203.

MacFarlane, P., Anderson, T., & McClintock, A.S. (2015). The Early Formation of the Working Alliance From the Client's Perspective: A Qualitative Study. *Psychotherapy, 52,* 363-372. DOI: 10.1037/a0038733

MacLeod, R., & Elliott, R. (2012). Emotion-focused therapy for social anxiety: A hermeneutic single-case efficacy design of a low-outcome case. *Counselling Psychology Review, 27*,7 –21.

MacLeod, R. & Elliott, R. (2014). Nondirective Person-centered therapy for social anxiety: a hermeneutic single-case efficacy design study of a good outcome case. *Person-Centered & Experiential Psychotherapies, 13,* 294-311.

Maling, M.S., Gurtman, M.B., & Howard, K.I. (1995). The response of interpersonal problems to varying doses of psychotherapy. *Psychotherapy Research, 5,* 63-75.

Margison, F.R., Barkham, M., Evans, C., McGrath, G., Mellor-Clark, J., Audin, K., & Connell, J. (2000). Measurement and psychotherapy: Evidence-based practice and practice-based evidence. *British Journal of Psychiatry, 177,* 123-130.

McElvaney, J. & Timulak, L. (2013). Clients' experience of therapy and its outcomes in 'good' and 'poor' outcome psychological therapy in a primary care setting: An exploratory study. *Counselling & Psychotherapy Research, 13,* 246-253. DOI: 10.1080/14733145.2012.761258

McLeod, J. (2001). An administratively created reality: some problems with the use of self-report questionnaire measures of adjustment in counselling/ psychotherapy outcome research. *Counselling & Psychotherapy Research, 1,* 215-226.

McLeod, J. (2010). *Case Study Research in Counselling and Psychotherapy*. London: Sage.

McLeod, J. (2011). *Qualitative Research in Counselling and Psychotherapy*. London: Sage.

Mearns, D. (1997). *Person-Centred Counselling Training.* London: Sage.

Mearns, D. (2003). *Developing Person-Centred Counselling.* London: Sage.

Mearns, D. & Thorne, B. (2013). *Person-Centred Counselling in Action*. London: Sage.

Mellor-Clark, J., Cross, S., Macdonald, J., & Skjulsvik, T. (2016). Leading Horses to Water: Lessons from a Decade of Helping Psychological Therapy Services Use Routine Outcome Measurement to Improve Practice. *Administration and Policy in Mental Health and Mental Health Services Research, 43,* 279-285.

Messick, S. (1995). Validity of Psychological Assessment: Validation of inferences from person's responses and performances as scientific inquiry into score meaning. *American Psychologist, 50,* 741-749.

Morrow-Bradley, C. & Elliott, R. (1986). Utilization of Psychotherapy Research by Practicing Psychotherapists. *American Psychologist, 41,* 188-197.

Murphy, D., Demetriou, E., & Joseph, S. (2015). A cross-sectional study to explore the mediating effect of intrinsic aspiration on the association between unconditional positive self-regard and posttraumatic growth. *Person-Centered & Experiential Psychotherapies, 14,* 201-213.

Murphy, D., Joseph, S., Demetriou, E., & Karimi-Mofrad, P. (2017). Unconditional Positive Self-Regard, Intrinsic Aspirations, and Authenticity: Pathways to Psychological Well-Being. *Journal of Humanistic Psychology,* 1-22. DOI: 10.1177/2F0022167816688314

Neff, K.D. (2003). The Development and Validation of a Scale to Measure Self-Compassion. *Self and Identity, 2,* 223-250.

Neff, K.D., Long, P., Knox, M.C., Davidson, O., Kuchar, A., Costigan, A., Williamson, Z., Rohleder, N., Tóth-Király, I., & Breines, J.G. (2018). The forest and the trees: Examining the association of self-compassion and its positive and negative components with psychological functioning. *Self and Identity, 17,* 627-645. DOI: 10.1080/15298868.2018.1436587

Neff, K.D., Whittaker, T.D., & Karl, A. (2017). Examining the Factor Structure of the Self-Compassion Scale in Four Distinct Populations: Is the Use of a Total Scale Score Justified? *Journal of Personality Assessment, 99,* 596-607.

Neville, B. (2013). Anxiously congruent: congruently anxious. *Person-Centered & Experiential Psychotherapies, 12,* 223-236.

Ogles, B.M. (2013). Measuring change in psychotherapy research. In M.J. Lambert (Ed.). *Bergin & Garfield's Handbook of Psychotherapy & Behavior Change* (pp.134-166). Hoboken, N.J.: Wiley.

Ollendick, T.H. (2014). Advances Toward Evidence-Based Practice: Where to From Here? *Behavior Therapy, 45,* 51-55.

Owen, J., Adelson, J., Budge, S., Wampold, B., Kopta, M., Minami, T., & Miller, S. (2015). Trajectories of change in psychotherapy. *Journal of Clinical Psychology, 71,* 817-827.

Pascual-Leone, A. (2018). How clients "change emotion with emotion": A programme of research on emotional processing. *Psychotherapy Research, 28,* 165-182. DOI: 10.1080/10503307.2017.1349350

Pascual-Leone, A. & Yeryomenko, N. (2017). The client "experiencing" scale as a predictor of treatment outcomes: A meta-analysis on psychotherapy process. *Psychotherapy Research, 27,* 653-665.

Pastor, D.A., & Beretvas, S.N. (2006). Longitudinal Rasch Modeling in the Context of Psychotherapy Outcomes Assessment. *Applied Psychological Measurement, 30,* 100-120.

Patterson, T.G. (2017). Assessing efficacy and effectiveness in person-centred therapy: challenges and opportunities. In S. Joseph (Ed.) *The Handbook of Person-Centred Therapy and Mental Health: Theory, Research and Practice* (pp.11-36). Monmouth: PCCS Books Ltd.

Patterson, T.G. & Joseph, S. (2006). Development of a self-report measure of unconditional positive self-regard. *Psychology and Psychotherapy: Theory, Research and Practice: Theory, Research and Practice, 79,* 557-570.

Philippot, P., Dethier, V., Baeyens, C., & Bouvard, M. (2018). Validation of the Self-Discrepancies Scale (S-DS). A tool to investigate the self in clinical and research settings. *Revue européenne de psychologie appliquée, 68,* 69–77.

Plake, B.S. & Wise, L.L. (2014). What is the role and importance of the revised AERA, APA, NCME Standards for Educational and Psychological Testing? *Educational Measurement: Issues and Practice, 33(4),* 4-12.

Proctor, C., Tweed, R. & Morris, D. (2016). The Rogerian Fully Functioning Person: A Positive Psychology Perspective. *Journal of Humanistic Psychology, 56,* 503-529.

Proctor, G. & Hayes, C. (2017). Counselling for Depression: a response to counselling education in the twenty-first century. Ethical conflicts for a counselling approach operating within a medicalised bureaucratic health service. *British Journal of Guidance & Counselling, 45*, 417-426.

Purton, C. (2013). Incongruence and "inner experience". *Person-Centered & Experiential Psychotherapies, 12*, 187-199.

Rennie, D.L. (1994). Clients' deference in psychotherapy. *Journal of Counselling Psychology, 41,* 427-437.

Rennie, D.L. (1998). *Person-Centred Counselling: An Experiential Approach.* London: Sage.

Rennie, D.L. (2000). Grounded Theory Methodology as Methodical Hermeneutics: Reconciling realism and relativism. *Theory & Psychology, 10,* 481-502.

Rolstad, S., Adler, J., & Rydén, A. (2011). Response Burden and Questionnaire Length: Is Shorter Better? A Review and Meta-analysis. *Value in Health, 14*, 1101-1108.

Rogers, C.R. (1942). *Counselling and Psychotherapy.* Boston MA: Houghton Mifflin.

Rogers, C.R. (1951). *Client-Centred Therapy*. London: Constable.

Rogers, C.R. (1957/1992). The Necessary and Sufficient Conditions of Therapeutic Personality Change. *Journal of Consulting and Clinical Psychology, 60,* 827-832.

Rogers, C.R. (1959). A theory of therapy, personality, and interpersonal relationships, as developed in the client-centered framework. In S. Koch (Ed.) *Psychology: A Study of a Science, Volume 3, Formulations of the Person and the Social Context* (pp.184-256). New York: McGraw-Hill.

Rogers, C.R. (1961/1967). *On Becoming A Person: A Therapist's View of Psychotherapy.* London: Constable.

Rogers, C.R. (1963). The concept of the fully functioning person. *Psychotherapy: Theory, Research & Practice, 1,* 17-26.

Rogers, C.R. & Dymond, R.F. (Eds.) (1954). *Psychotherapy and Personality Change: Coordinated research studies in the client-centered approach.* Chicago, IL: The University of Chicago Press.

Sanders, P. (2017). Principled and strategic opposition to the medicalization of distress and all of its apparatus. In S. Joseph (Ed.) *The Handbook of Person-Centred Therapy and Mental Health: Theory, Research and Practice* (pp.11-36). Monmouth: PCCS Books Ltd.

Sanders, T., & Freire, E. (2008). Researching the relationship: Developing the Therapeutic Relationship Scale. Paper presented at conference of the World Association for Person-Centred and Experiential Psychotherapy and Counselling, Norwich UK.

Schmid, P.F. (2001). Authenticity: The person as his or her own author. Dialogical and ethical perspectives on therapy as an encounter relationship. And Beyond. In G. Wyatt (Ed.) *Congruence* (pp.213-228). Ross-on-Wye: PCCS Books.

Seeman, J. (2008). *Psychotherapy and the Fully Functioning Person*. Bloomington, IN: AuthorHouse.

Sheldon, K.M. (1995). Creativity and Self-Determination in Personality. *Creativity Research Journal, 8,* 25-36.

Sheldon, K.M., Ryan, R., & Reis, H.T. (1996). What makes for a good day? Competence and autonomy in the day and the person. *Personality and Social Psychology Bulletin, 22,* 1270-1279.

Shostrom, E.L. (1964). An inventory for the measurement of self-actualization. *Educational and Psychological Measurement, 24,* 207-218.

Slaney, K.L. (2017). *Validating Psychological Constructs: Historical, Philosophical and Practical Dimensions.* London: Palgrave MacMillan.

Slaney, K.L. & Maraun, M.D. (2008). A Proposed Framework for Conducting Data-Based Test Analysis. *Psychological Methods, 13,* 376-390. DOI: 10.1037/a0014269.

Slaney, K.L., Storey, J.E., & Barnes, J. (2011a). Is My Test Valid? Guidelines for the Practicing Psychologist for Evaluating the Psychometric Properties of Measures. *International Journal of Forensic Mental Health, 10,* 261-283. DOI: 10.1080/14999013.2011.627086

Slaney, K.L., Storey, J.E., & Barnes, J. (2011b). When "Good Enough" Is Just Not Good Enough: Response to Holden and Marjanovic. *International Journal of Forensic Mental Health, 10,* 290-294. DOI: 10.1080/14999013.2011.629716

Smith, B.W., Guzman, A. & Erikson, K. (2018). The Unconditional Self-Kindness Scale: Assessing the Ability to Respond with Kindness to Threats to the Self. *Mindfulness, 9,* 1713-1722.

Spector, P.E., Van Katwyk, P.T., Brannick, M.T. & Chen, P.Y. (1997). When Two Factors Don't Reflect Two Constructs: How Item Characteristics Can Produce Artifactual Factors. *Journal of Management, 23,* 659-677.

Stephen, S. & Elliott, R. (2011). Developing the Adjudicated Case Study Method. *Pragmatic Case Studies in Psychotherapy, 7*, 230-241.

Stephen, S., Elliott, R., & MacLeod, R. (2011). Person-centred therapy with a client experiencing social anxiety difficulties: A hermeneutic single case efficacy design. *Counselling & Psychotherapy Research, 11*, 55-66.

Stevens, F.L. (2017). Authenticity: a mediator in the relationship between attachment style and affective functioning. *Counselling Psychology Quarterly, 30,* 392-414.

Stiles, W.B., Barkham, M., Connell, J. & Mellor-Clark, J. (2008). Responsive Regulation of Treatment Duration in Routine Practice in United Kingdom

Primary Care Settings: Replication in a Larger Sample. *Journal of Consulting and Clinical Psychology, 76,* 298-305.

Stiles, W. B., Barkham, M., Twigg, E., Mellor-Clark, J. & Cooper, M. (2006). Effectiveness of cognitive-behavioural, person-centred and psychodynamic therapies as practised in UK National Health Service settings. *Psychological Medicine, 36,* 555-566.

Stone, C. & Elliott, R. (2011). Clients' experience of research within a research clinic setting. *Counselling Psychology Review, 26,* 71-86.

Strahan, R. & Gerbasi, K C. (1972). Short, homogeneous versions of the Marlowe-Crowne Social Desirability Scale. *Journal of Clinical Psychology, 28*, 191-193.

Stulz, N., Lutz, W., Leach, C., Lucock, M., & Barkham, M. (2007). Shapes of Early Change in Psychotherapy Under Routine Outpatient Conditions. *Journal of Consulting & Clinical Psychology, 75,* 864-874.

Surviving Work (2019). *The IAPT Survey Results*. Retrieved from: https://survivingwork.org/5241-2/

Tangney, J.P., Niedenthal, P.M., Covert, M.V., & Barlow, D.H. (1998). Are Shame and Guilt Related to Distinct Self-Discrepancies? A Test of Higgins's (1987) Hypotheses. *Journal of Personality and Social Psychology, 75,* 256-268.

Tashiro, C. (2011). *Evaluating client change in person-centred/experiential therapies in a research clinic setting: Client outcomes and analysis of outcome measures*. Unpublished MSc dissertation, University of Strathclyde.

Timulak, L. (2007). Identifying core categories of client-identified impact of helpful events in psychotherapy: A qualitative meta-analysis. *Psychotherapy Research, 17,* 305-314.

Timulak, L. (2009). Meta-analysis of qualitative studies: A tool for reviewing qualitative findings in psychotherapy. *Psychotherapy Research, 19*, 591-600.

Tolan, J. (2017). *Skills in Person-Centred Counselling & Psychotherapy.* London: Sage.

Tosi, D.J. & Lindamood, C.A. (1975). The Measurement of Self-Actualization: A Critical Review of the Personal Orientation Inventory. *Journal of Personality Assessment, 39,* 215-224.

Toukmanian, S.G., Jadaa, D-A., Armstrong, M.S. (2010). Change Processes in Clients' Self-Perceptions in Experiential Psychotherapy. *Person-Centered & Experiential Psychotherapies, 9,* 37-51.

Traynor, W. (2019). *Person-centred therapy and pre-therapy for people who hear voices, have unusual experiences or psychotic processes: Practitioner and client perceptions of helpful and unhelpful practice and perceived client changes*. Unpublished PhD thesis, University of Strathclyde.

Truijens, F.L. (2017). Do the numbers speak for themselves? A critical analysis of procedural objectivity in psychotherapeutic efficacy research. *Synthese, 194,* 4721-4740. DOI 10.1007/s11229-016-1188-8

Truijens, F.L., Desmet, M., De Coster, E., Uyttenhove, H., Deeren, B., & Meganck, R. (2019). When quantitative measures become a qualitative storybook: A phenomenological case analysis of validity and performativity of questionnaire administration in psychotherapy research. *Qualitative Research in Psychology.* DOI:10.1080/14780887.2019.1579287

Tudor, K. (2008a). Verbal being: From being human to human being. In B.E. Levitt (Ed.) *Reflections on Human Potential: Bridging the person-centered approach and positive psychology* (pp.68-83). Ross-on-Wye: PCCS Books Ltd.

Tudor, K. (2008b). Psychological health: Autonomy and homonomy. In B.E. Levitt (Ed.) *Reflections on Human Potential: Bridging the person-centered approach and positive psychology* (pp.161-174). Ross-on-Wye: PCCS Books Ltd.

Tudor, L.E., Keemar, K., Tudor, K., Valentine, J. & Worrall, M. (2004). *The Person-Centred Approach: A Contemporary Introduction.* Basingstoke: Palgrave MacMillan.

Tzeng, O.C.S., Maxey, W.A., Fortier, R. & Landis, D. (1985). Construct Evaluation of the Tennessee Self Concept Scale. *Educational and Psychological Measurement, 45,* 63-78.

Vaidya, D. (2013). Re-visioning Rogers' second condition – anxiety as the face of ontological incongruence and the basis for the principle of non-directivity in PCT therapy. *Person-Centered & Experiential Psychotherapies, 12,* 209-222.

Veale, D., Kinderman, P., Riley, S. & Lambrou, C. (2003). Self-discrepancy in bod dysmorphic disorder. *British Journal of Clinical Psychology, 42,* 157-169.

Walker, A.M., Rablen, R.A., & Rogers, C.R. (1960). Development of a scale to measure process changes in psychotherapy. *Journal of Clinical Psychology, 16,* 79-85.

Warner, M.S. (2000). Person-centered psychotherapy: One nation, many tribes. *The Person-Centered Journal, 7,* 28-39.

Warner, M.S. (2009). Defense or Actualization? Reconsidering the Role of Processing, Self and Agency within Rogers' Theory of Personality. *Person-Centered & Experiential Psychotherapies, 8*, 109-126.

Warner, M.S. (2017). A person-centred view of human nature, wellness and psychopathology. In S. Joseph (Ed.). *The Handbook of Person-Centred Therapy and Mental Health: Theory, Research and Practice* (pp.92-115). Monmouth: PCCS Books Ltd.

Watson, J. (Ed.) (2019). *Drop the Disorder: Challenging the culture of psychiatric diagnosis.* Monmouth: PCCS Books Ltd.

Watson, J., Goldman, R., & Greenberg, L.S. (2007). *Case Studies in Emotion-Focused Therapy for Depression*. Washington, DC: American Psychological Association.

Watson, J.C. & Lilova, S. (2009). Testing the Reliability and Validity of the Scales for Experiencing Emotion with a Canadian Sample. *Person-Centered & Experiential Psychotherapies, 8,* 189-207.

Watson, J.C. & Watson, N. (2010). Operationalizing Incongruence: Measures of self-discrepancy and affect regulation. In M. Cooper, J.C. Watson & D. Hölldampf (Eds.) *Person-Centered and Experiential Therapies Work* (pp.164-187). Ross-on-Wye: PCCS Books Ltd.

Watson, N., Bryan, B.C., & Thrash, T.M. (2010). Self-Discrepancy: Comparisons of the Psychometric Properties of Three Instruments. *Psychological Assessment, 22,* 878-892.

Watson, N., Bryan, B.C., & Thrash, T.M. (2014). Change in Self-Discrepancy, Anxiety, and Depression in Individual Therapy. *Psychotherapy, 51,* 525-534.

Watson, N., Bryan, B.C., & Thrash, T.M. (2016). Self-Discrepancy: Long-Term Test–Retest Reliability and Test–Criterion Predictive Validity. *Psychological Assessment, 28,* 59-69.

Weiss, A.S. (1987). Shostrom's Personal Orientation Inventory: Arguments against its basic validity. *Personality and Individual Differences, 8,* 895-903.

Werbart, A., Annevall, A., & Hillblom, J. (2019). Successful and less successful psychotherapies compared: Three therapists and their six contrasting cases. *Frontiers in Psychology, 10,* art.816. DOI: 10.3389/fpsyg.2019.00816

Werbart, A., von Below, C., Engqvist, K. & Lind, S. (2018). "It was like having half of the patient in therapy": Therapists of nonimproved patients looking back on their work. *Psychotherapy Research,* DOI: 10.1080/10503307.2018.1453621

Widdowson, M.D.J. (2013). *The Process and Outcome of Transactional Analysis Psychotherapy for the Treatment of Depression: An Adjudicated Case Series*. Unpublished PhD dissertation, University of Leicester.

Wiggins, S., Elliott, R., & Cooper, M. (2012). The prevalence and characteristics of relational depth events in psychotherapy. *Psychotherapy Research, 22*, 139-158.

Wood, A.M., Linley, P.A., Maltby, J., Baliousis, M., & Joseph, S. (2008). The Authentic Personality: A Theoretical and Empirical Conceptualization and the

Development of the Authenticity Scale. *Journal of Counseling Psychology, 55,* 385-399.

Wyatt, G. (Ed.) (2001). *Congruence*. Ross-on-Wye: PCCS Book

Zech, E., Brison, C., Elliott, R., Rodgers, B., & Cornelius-White, J.H.D. (2018). Measuring Rogers' conception of personality development: validation of the Strathclyde Inventory-French version. *Person-Centered & Experiential Psychotherapies, 17*, 160-184.

Zhou, Y. (2019). A Mixed Methods Model of Scale Development and Validation Analysis. *Measurement: Interdisciplinary Research and Perspectives, 17,* 38-47.

Appendix A: Conference presentations

Stephen, S. (2016). *The Strathclyde Inventory: Next steps in the development of a person-centered outcome measure*. Paper presented at the 12th PCE Conference, New York City, U.S.A.

Stephen, S. & Elliott, R. (2017). *The Strathclyde Inventory: Measuring congruence as an outcome of therapy.* Paper presented at the 23rd BACP Research Conference, Chester, UK.

Stephen, S. & Elliott, R. (2018). *Investigating patterns of change in client congruence during person-centered therapy*. Paper presented at 13th PCE Conference, Vienna, Austria.

Stephen, S., Bell, L., Kahn, M., Love, R., Mackintosh, H., Martin, M., Moran, R., Price, E. & Whitehead, B. (2019a). *Deterioration as an outcome of therapy: A systematic case study group investigation*. Symposium presented at the 25th BACP Research Conference, Belfast, UK.

Stephen, S., Bell, L., Kahn, M., Love, R., Mackintosh, H., Martin, M., Moran, R., Price, E. & Whitehead, B. (2019b). *What features of therapy may contribute to a client's deterioration? Meta-synthesis of a systematic case study series.* Paper presented at the 25th BACP Research Conference, Belfast, UK.

Appendix B: Ethical approval of research clinic generic framework (renewed April 2018)

Approval: UEC17/73 Elliot: Generic Framework Application – Practice-Based Psychotherapy Research Clinic Protocol, Phase 2 (RESUBMISSION FROM NOVEMBER 2017)

Ethics
Sent: 10 April 2018 13:25
To: Robert Elliott
Cc: Susan Stephen; Lorna Carrick; Ethics

Dear Robert

ETHICAL AND SPONSORSHIP APPROVAL
UEC17/73 Elliot: Generic Framework Application – Practice-Based Psychotherapy Research Clinic Protocol, Phase 2 (RESUBMISSION FROM NOVEMBER 2017)

I can confirm that the University Ethics Committee (UEC) has approved this protocol and appropriate insurance cover and sponsorship have now also been confirmed.

I would remind you that the UEC must be informed of any changes you plan to make to the research project, so that it has the opportunity to consider them. Any change of staffing within the research team should be reported to UEC.

The UEC would also expect you to report back on the progress and outcome of your project, with an account of anything which may prompt ethical questions for any similar future project and with anything else that you feel the Committee should know.

Any adverse event that occurs during an investigation must be reported as quickly as possible to UEC and, within the required time frame, to any appropriate external agency.

The University agrees to act as sponsor of the above mentioned project subject to the following conditions:

1. That the project obtains/has and continues to have University/Departmental Ethics Committee approval.

2. That the project is carried out according to the project protocol.

3. That the project continues to be covered by the University's insurance cover.

4. That the Director of Research and Knowledge Exchange Services is immediately notified of any change to the project protocol or circumstances which may affect the University's risk assessment of the project.

5. That the project starts within 12 months of the date of this letter.

As sponsor of the project the University has responsibilities under the Scottish Executive's Research Governance Framework for Health and Community Care. You should ensure you are aware of those responsibilities and that the project is carried out according to the Research Governance Framework.

On behalf of the Committee, I wish you success with this project.

Kind regards
Angelique

Angelique Laverty

Appendix C: Strathclyde Inventory – 16 items (SI-16)

Please read each statement below and think how often you sense it has been true for you DURING THE **LAST MONTH**. Then mark the box that is closest to this. There are no right or wrong answers – it is only important what is true for you individually.

OVER THE LAST MONTH	Never	Only Occasionally	Sometimes	Often	All or most of the time
1. I have been able to be spontaneous	☐0	☐1	☐2	☐3	☐4
2. I have condemned myself for my attitudes or behaviour	☐4	☐3	☐2	☐1	☐0
3. I have tried to be what others think I should be	☐4	☐3	☐2	☐1	☐0
4. I have trusted my own reactions to situations	☐0	☐1	☐2	☐3	☐4
5. I have experienced very satisfying personal relationships	☐0	☐1	☐2	☐3	☐4
6. I have felt afraid of my emotional reactions	☐4	☐3	☐2	☐1	☐0
7. I have looked to others for approval or disapproval	☐4	☐3	☐2	☐1	☐0
8. I have expressed myself in my own unique way	☐0	☐1	☐2	☐3	☐4
9. I have found myself "on guard" when relating with others	☐4	☐3	☐2	☐1	☐0
10. I have made choices based on my own internal sense of what is right	☐0	☐1	☐2	☐3	☐4
11. I have listened sensitively to myself	☐0	☐1	☐2	☐3	☐4
12. I have lived fully in each new moment	☐0	☐1	☐2	☐3	☐4
13. I have hidden some elements of myself behind a "mask"	☐4	☐3	☐2	☐1	☐0
14. I have felt true to myself	☐0	☐1	☐2	☐3	☐4
15. I have been able to resolve conflicts within myself	☐0	☐1	☐2	☐3	☐4
16. I have felt it is all right to be the kind of person I am	☐0	☐1	☐2	☐3	☐4

Appendix D: Strathclyde Inventory – 20 items (SI-20)

Please read each statement below and think how often you sense it has been true for you DURING THE **LAST MONTH**. Then mark the box that is closest to this. There are no right or wrong answers – it is only important what is true for you individually.

OVER THE LAST MONTH	Never	Only Occasionally	Sometimes	Often	All or most of the time
1. I have been able to be spontaneous	☐ 0	☐ 1	☐ 2	☐ 3	☐ 4
2. I have condemned myself for my attitudes or behaviour	☐ 4	☐ 3	☐ 2	☐ 1	☐ 0
3. I have tried to be what others think I should be	☐ 4	☐ 3	☐ 2	☐ 1	☐ 0
4. I have trusted my own reactions to situations	☐ 0	☐ 1	☐ 2	☐ 3	☐ 4
5. I have experienced very satisfying personal relationships	☐ 0	☐ 1	☐ 2	☐ 3	☐ 4
6. I have felt afraid of my emotional reactions	☐ 4	☐ 3	☐ 2	☐ 1	☐ 0
7. I have looked to others for approval or disapproval	☐ 4	☐ 3	☐ 2	☐ 1	☐ 0
8. I have expressed myself in my own unique way	☐ 0	☐ 1	☐ 2	☐ 3	☐ 4
9. I have found myself "on guard" when relating with others	☐ 4	☐ 3	☐ 2	☐ 1	☐ 0
10. I have made choices based on my own internal sense of what is right	☐ 0	☐ 1	☐ 2	☐ 3	☐ 4
11. I have listened sensitively to myself	☐ 0	☐ 1	☐ 2	☐ 3	☐ 4
12. I have felt myself doing things that were out of my control	☐ 4	☐ 3	☐ 2	☐ 1	☐ 0
13. I have lived fully in each new moment	☐ 0	☐ 1	☐ 2	☐ 3	☐ 4
14. I have been aware of my feelings	☐ 0	☐ 1	☐ 2	☐ 3	☐ 4
15. I have hidden some elements of myself behind a "mask"	☐ 4	☐ 3	☐ 2	☐ 1	☐ 0
16. I have felt true to myself	☐ 0	☐ 1	☐ 2	☐ 3	☐ 4
17. I have been able to resolve conflicts within myself	☐ 0	☐ 1	☐ 2	☐ 3	☐ 4
18. I have felt myself doing things that are out of character for me	☐ 4	☐ 3	☐ 2	☐ 1	☐ 0
19. I have accepted my feelings	☐ 0	☐ 1	☐ 2	☐ 3	☐ 4
20. I have felt it is all right to be the kind of person I am	☐ 0	☐ 1	☐ 2	☐ 3	☐ 4

Appendix E: Strathclyde Inventory – 12 items (SI-12)

Please read each statement below and think how often you sense it has been true for you DURING THE **LAST MONTH**. Then mark the box that is closest to this. There are no right or wrong answers – it is only important what is true for you individually.

OVER THE LAST MONTH	Never	Only Occasionally	Sometimes	Often	All or most of the time
1. I have been able to be spontaneous	☐ 0	☐ 1	☐ 2	☐ 3	☐ 4
2. I have condemned myself for my attitudes or behaviour	☐ 4	☐ 3	☐ 2	☐ 1	☐ 0
3. I have tried to be what others think I should be	☐ 4	☐ 3	☐ 2	☐ 1	☐ 0
4. I have trusted my own reactions to situations	☐ 0	☐ 1	☐ 2	☐ 3	☐ 4
5. I have found myself "on guard" when relating with others	☐ 4	☐ 3	☐ 2	☐ 1	☐ 0
6. I have listened sensitively to myself	☐ 0	☐ 1	☐ 2	☐ 3	☐ 4
7. I have felt myself doing things that were out of my control	☐ 4	☐ 3	☐ 2	☐ 1	☐ 0
8. I have lived fully in each new moment	☐ 0	☐ 1	☐ 2	☐ 3	☐ 4
9. I have been aware of my feelings	☐ 0	☐ 1	☐ 2	☐ 3	☐ 4
10. I have hidden some elements of myself behind a "mask"	☐ 4	☐ 3	☐ 2	☐ 1	☐ 0
11. I have felt myself doing things that are out of character for me	☐ 4	☐ 3	☐ 2	☐ 1	☐ 0
12. I have felt it is all right to be the kind of person I am	☐ 0	☐ 1	☐ 2	☐ 3	☐ 4

Appendix F: MSc Dissertations

Bell, L. (2018). *A good 'person-centred outcome' based on the Strathclyde Inventory: A hermeneutic single case efficacy design.* Unpublished MSc dissertation, University of Strathclyde.

Khan, M. (2018). *Person-Centred Therapy efficacy with client experiencing Social Anxiety: A hermeneutic single case efficacy design study*. Unpublished MSc dissertation, University of Strathclyde.

Love, R. (2018). *Deterioration over the course of person-centred therapy: A Hermeneutic Single Case Efficacy Design*. Unpublished MSc dissertation, University of Strathclyde.

MacIntosh, H. (2018). *The mediating and moderating factors that affect outcome in person-centred therapy for clients who experience a reliable deterioration on the Strathclyde Inventory; a hermeneutic single case efficacy design* (HSCED). Unpublished MSc dissertation, University of Strathclyde.

Martin, M. (2018). *Person-centred therapy with a client experiencing depression, psychosis, & anxiety: a hermeneutic single case efficacy design study.* Unpublished MSc dissertation, University of Strathclyde.

Moran, R, (2018). *Hermeneutic Single-Case Efficacy Design: The impact of person-centred therapy on a negative outcome case*. Unpublished MSc dissertation, University of Strathclyde.

Price, E. (2018). *A Hermeneutic single case efficacy design study of a client indicating post-therapy deterioration on the Strathclyde Inventory.* Unpublished MSc dissertation, University of Strathclyde.

Whitehead, B.T. (2018). *Person-centered therapy with a client with Asperger Syndrome, an Autism Spectrum Disorder (ASD): A hermeneutic single case efficacy design study*. Unpublished MSc dissertation, University of Strathclyde.

Appendix G: Consent Form

Strathclyde Centre for Counselling and Psychotherapy
Suite D303 David Stow Building, Jordanhill Campus
University of Strathclyde Counselling Unit
76 Southbrae Drive, Glasgow G13 1PP
Email: enquiries@strathclydetherapy.com
Phone: 0844 586 4560

PRACTICE-BASED PSYCHOTHERAPY RESEARCH

CONSENT FORM (v5; 09/11)

		Please initial box
1.	I confirm that I have read and understand the information sheet dated 09/2011 (v5) for the above study. I have had the opportunity to consider the information, ask questions, and have these answered satisfactorily.	☐
2.	I understand that my participation is voluntary and that I am free to withdraw at any time without giving any reason, without my legal rights being affected.	☐
3.	I understand that relevant data collected during the study may be used by members of the research team at the University of Strathclyde. I understand that I will be asked separately about the use of the recordings of my counselling sessions and research interviews as detailed in the Release of Recordings form dated 09/2011 (v5).	☐
4.	I confirm that I am aged 18 or over and that I am aware of what my participation involves and any potential risks.	☐
5.	I agree to take part in this study	☐

Name of participant	Date	Signature
Name of researcher /witness	Date	Signature

Appendix H: Release of Recordings Form

Strathclyde Centre for Counselling and Psychotherapy
Suite D303 David Stow Building, Jordanhill Campus
University of Strathclyde Counselling Unit
76 Southbrae Drive, Glasgow G13 1PP
Email: enquiries@strathclydetherapy.com
Phone: 0844 586 4560

PRACTICE-BASED PSYCHOTHERAPY RESEARCH

RELEASE OF RECORDINGS CONSENT FORM (v5; 09/2011)

Once you have finished your counselling, we would like your permission to use the recordings of your research interviews and therapy sessions to help us understand how therapy works. Below are some of the possible situations in which we would like to use these recordings, if you are willing to give us permission to do so.

For each of the situations described below, please indicate whether you agree to this use or not. Please don't agree to anything you feel uncomfortable with. We are asking you to review this form after ten sessions and again at the end of counselling so that you can make changes if you wish to. Please feel free to discuss this with your counsellor and to negotiate with the research assistant about any of these possible uses.

		Please circle one	Please initial box
1.	After counselling is over, I am willing for my counsellor to read the questionnaires and listen to what I said in the research interviews.	NO YES	
2.	I am willing for the video and audio recordings of my sessions to be used for training other therapists or counsellors in the present project, for a period of at least 5 years.	NO YES	
3.	I am willing for the video and audio recordings of my counselling sessions and research interviews to be used for training other postgraduate level students or other mental health professionals, for a period of at least 5 years or as long as there is a specific use identified by the Chief Investigator or research team.	NO YES	
4.	I am willing for the professional members (the investigators, research associates, postgraduate	NO YES	

counselling students, and professional consultants) of the research team to analyse the recordings for the purpose of developing and evaluating Person-Centred and Experiential psychotherapies.

5. I am willing for brief excerpts from my counselling sessions and research interviews to be presented at scientific meetings or in scientific publications in order to better understand what the therapeutic process is like for clients. I am willing for these excerpts to take the form of: (please cross out any which you wish to exclude):
    - anonymous transcripts of counselling sessions
    - audio recordings of counselling sessions
    - video recordings of counselling sessions
    - anonymous transcripts of research interviews
    - audio/video recordings of research interviews

NO YES ☐

6. I am willing for the information that I have given in my research questionnaires and interviews, as well as extracts from therapy sessions, to be analysed and presented as a systematic single case study.

NO YES ☐

7. I am willing for research teams at other Universities within the European Union to analyse data from my counselling as long as they are monitored by the Chief Investigator and pledge to protect my identity. This permission includes (please cross any which you wish to exclude):
    - questionnaire data
    - anonymous transcripts of counselling sessions
    - audio recordings of counselling sessions
    - video recordings of counselling sessions
    - anonymous transcripts of research interviews
    - audio/video recordings of research interviews

NO YES ☐

8. I am willing for research teams at Universities outside the European Union, which are not covered by the Data Protection Act, to analyse data from my counselling as long as they are monitored by the Chief Investigator and pledge to protect my identity. This permission includes (please cross any which you wish to exclude):
    - questionnaire data
    - anonymous transcripts of counselling sessions
    - audio recordings of counselling sessions
    - video recordings of counselling sessions
    - anonymous transcripts of research interviews
    - audio/video recordings of research interviews

NO YES ☐

9. I am willing to be contacted if any additional use of the recordings or other data is requested, including reviewing or commenting on systematic single case study reports.

NO
YES

Please indicate specific identifying information which should be edited from the recordings (e.g. personal names, place names, places of employment or schools):

Please indicate a permanent address and phone number or email address at which you may be contacted:

I understand that, by responding to the above items and signing below, I have given my permission for the video and audio recordings and other data from my sessions and interviews to be used in the manner I have specified.

_____	_____	_____
Name of participant	Date	Signature
_____	_____	_____
Name of researcher /witness	Date	Signature

Appendix I: HSCED Adjudication Panel Decision Pro Forma (2018 Version)

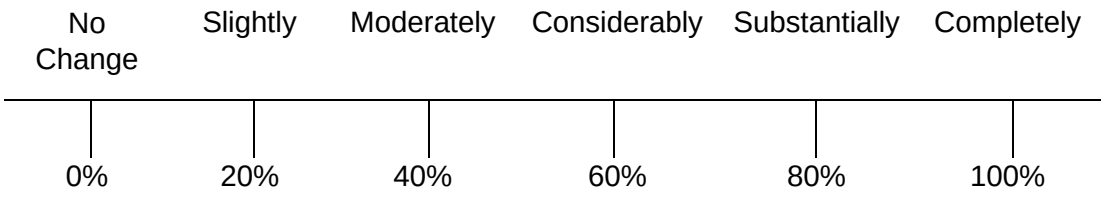
Completing the adjudication process

It is the aim of the Adjudication Panel to reach a joint decision in answer to each question.

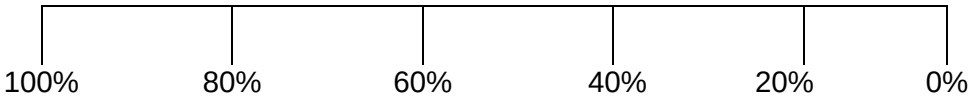
Please report your answers on the scales provided.

When reporting the reasons for each decision, please use whatever space you need in order to give a full response.

1. To what extent do you think the Client changed (improved/deteriorated) over the course of therapy?

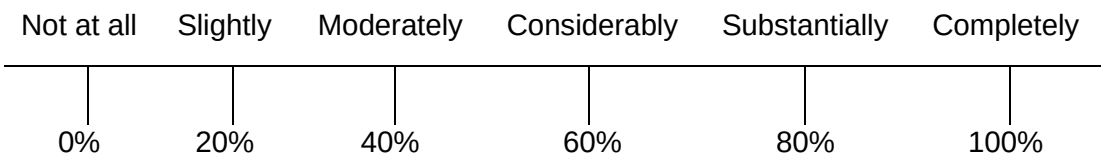


1a. How confident are you that the Client showed at least "substantial" change (improvement/deterioration) over the course of therapy?

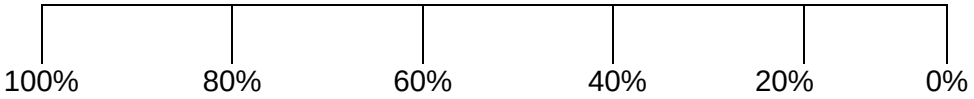


1b. What evidence presented in the affirmative and sceptic cases mattered most to you in reaching this conclusion? How did you make use of this evidence?

2. To what extent do you think that the Client's change (improvement/ deterioration) was due to the therapy?



2a. How confident are you that the Client's change was at least "substantially" due to therapy?



2b. What evidence presented in the affirmative and sceptic cases mattered most to you in reaching this conclusion? How did you make use of this evidence?

3. Which therapy processes (mediator factors) do you feel were helpful or hindering to the client?

4. Which characteristics and/or personal resources of the client (moderator factors) do you feel helped or hindered them to make best use of their therapy?

Appendix J: Overview of Study 3 Analysis

Notes. ↑ = results for improvers; ↓ = results for deteriorators; **bold** = name of domain, category or first level sub-category; results in greyscale = categories (dark grey), frequency = general or typical (light grey); * = difference of at least two frequency counts.

1	**IN THE BEGINNING**	N	↑	↓
1.1	**Client concerns**	8	4	4
1.1.1	**Self in the world**	8	4	4
1.1.1.1	Expressions of anxiety	6	4*	2
1.1.1.1.1	Stress / distress	2	1	1
1.1.1.1.2	Social anxiety	1	1	1
1.1.1.1.3	Existential anxiety	1	0	1
1.1.1.1.4	Panic	1	1	0
1.1.1.1.5	Psychosis	1	1	0
1.1.1.2	Relationship with others	6	2	4*
1.1.1.2.1	Feelings of isolation	2	1	1
1.1.1.2.2	Tendency to take on other people's issues	1	1	0
1.1.1.3	Contextual difficulties	3	1	2
1.1.1.3.1	Work-related problems	2	1	1
1.1.1.3.2	Adapting to life in UK	1	0	1
1.1.2	**Relationship with self**	6	2	4*
1.1.2.1	Depression	4	1	3*
1.1.2.2	Low self-esteem	4	1	3*
1.2	**Readiness to begin therapy**	8	4	4
1.2.1	**Motivated**	5	4	3
1.2.1.1	Determination	5	3	2
1.2.1.1.1	Stubborn	2	1	1
1.2.1.2	Expectations	4	2	2
1.2.1.3	Hope	3	2	1
1.2.2	**Open to experience**	6	4*	2
1.2.2.1	Able to be reflective	5	3	2
1.2.2.2	Willing to be open	3	2	1
1.2.2.3	Prepared to go to 'unhappy places'	2	1	1
1.2.3	**Prepared**	4	3*	1
1.2.3.1	Ready to get started	2	2	0
1.2.3.2	Proactive	2	2	0
1.2.3.2.1	Using other therapeutic services	2	2	0
1.2.3.4	Knew what wanted to work on	2	1	1
1.3	**Potential challenges to readiness**	6	3	3
1.3.1	**Doubt / potential for doubt**	4	1	3*
1.3.1.1	Expectation that would feel more vulnerable	1	0	1
1.3.1.2	'Wrong' expectations	1	0	1
1.3.2	**Difficulties in processing emotions and experience**	4	2	2

1.3.2.1	Difficulty in accessing and expressing emotions / experience	3	1	2
1.3.2.2	Difficulty understanding own experience	2	1	1
1.3.2.3	Wanting to avoid difficult emotions	2	1	1
1.3.3	**Tendency toward rigidity**	2	1	1
1.3.3.1	Perfectionism	1	1	0
1.3.3.2	Unwillingness to consider / accept another person's point of view	1	0	1
1.3.4	**Lack of supportive relationships**	2	0	2

2.	**IN THE PROCESS**	N	↑	↓
2.1	**Therapeutic climate**	8	4	4
2.1.1	**Safe**	6	3	3
2.1.1.1	Able to express feelings without upsetting someone else	2	1	1
2.1.1.2	Able to be real / authentic	2	1	1
2.1.1.3	Able to experience relational / emotional depth	1	1	0
2.1.1.4	Able to disclose / show vulnerability	2	1	1
2.1.2	**Warmth / connection**	6	3	3
2.1.3	**Supportive**	6	3	3
2.1.3.1	Space to process	2	1	1
2.1.3.2	Opportunity to practise communication skills	2	2	0
2.1.3.3	Support between sessions	2	0	2
2.1.4	**Collaborative**	4	2	2
2.2	**Therapist in the process**	8	4	4
2.2.1	**Facilitative therapist processes**	8	4	4
2.2.1.1	Accepting / understanding / validating client	8	4	4
2.2.1.1.1	Working within client's frame of reference	5	3	2
2.2.1.1.1.1	Using metaphor	2	2	0
2.2.1.1.1.2	Reflecting words back	1	1	0
2.2.1.1.1.3	Reflecting feeling	1	0	1
2.2.1.1.2	Validating/affirming client's experience	4	1	3*
2.2.1.1.3	Helping client feel comfortable	3	1	2
2.2.1.2	Offering opportunities	5	3	2
2.2.1.2.1	To ground / slow down client	3	2	1
2.2.1.2.2	To develop process	3	2	1
2.2.1.2.3	To engage client	2	1	1
2.2.1.2.4	To challenge client	2	1	1
2.2.1.3	Being transparent	5	3	2
2.2.1.3.1	Responding to client's questions	3	1	2
2.2.1.3.2	Discussing process	4	2	2
2.2.1.3.3	Noticing / validating changes	2	2	0
2.2.1.3.4	Expressing concern for client	1	1	0
2.2.2	**Interfering therapist processes**	4	1	3*
2.2.2.1	Not working in client's frame of reference / acting in a controlling /directive manner	3	0	3*
2.2.2.2	Therapist doubts	3	1	2
2.2.2.2.1	Unsure how/if helping client	2	1	1

2.2.2.2.2	Therapist at a loss to know what to do	2	0	2
2.2.2.2.3	Querying their connection with client	2	1	1
2.2.2.3	Inconsistent approach	1	0	1
2.3	**Client in the process**	8	4	4
2.3.1	**Facilitative client processes**	8	4	4
2.3.1.1	Client commitment	7	4	3
2.3.1.1.1	Engagement in therapeutic process	7	4	3
2.3.1.1.2	Integrating therapy and life	6	4*	2
2.3.1.1.2.1	Using therapy to connect with / work on life problem	6	4*	2
2.3.1.1.2.2	Using therapy to gain distance / perspective on life problem	2	2	0
2.3.1.1.2.3	Reflecting on progress in life problem	1	1	0
2.3.1.1.3	Working really hard in and out of therapy room	4	4*	0
2.3.1.1.3.1	Self-help / self-initiated efforts	2	2	0
2.3.1.2	Experiencing feelings (opening up / allowing self-awareness)	6	4*	2
2.3.1.2.1	Realising feelings	3	3*	0
2.3.1.2.2	Working with / resolving / releasing stuck feelings	2	2	0
2.3.1.2.2.1	Letting go	2	2	0
2.3.1.2.2.2	Letting in (inner child)	2	2	0
2.3.1.2.3	Greater general awareness of feelings	2	2	0
2.3.1.2.4	Valuing feelings	1	1	0
2.3.1.3	Working through complex situations	5	3	2
2.3.1.3.1	Exploring	5	3	2
2.3.1.3.1.1	Sorting thoughts more clearly	4	3*	1
2.3.1.3.1.2	Articulating thoughts	3	1	2
2.3.1.3.1.3	Identifying main issues	2	1	1
2.3.1.3.1.4	Looking at issues in depth	2	1	1
2.3.1.3.2	Making connections / understanding why things happened	3	3*	0
2.3.1.3.2.1	Recognising what issues are in the past	1	1	0
2.3.1.3.2.2	See current circumstances in larger context	1	1	0
2.3.2	**Interfering client processes**	4	0	4*
2.3.2.1	Discomfort with process	4	0	4*
2.3.2.1.1	Difficulty opening up	4	0	4*
2.3.2.1.1.1	Difficult and painful	3	0	3*
2.3.2.1.1.1.1	Pressure to engage?	3	0	3*
2.3.2.1.1.2	Feeling exposed	2	0	2
2.3.2.1.1.3	Concerned about upsetting therapist	1	0	1
2.3.2.1.2	Struggling to find direction	3	0	3*
2.3.2.1.2.1	Struggling to know what to say	1	0	1
2.3.2.1.2.2	Can feel like a chore to find useful direction for work	1	0	1
2.3.2.1.2.3	Coasting	1	0	1
2.3.2.1.3	Feeling that they had to change / do therapy differently / felt pressure from therapist	2	0	2

2.3.2.1.4	Feeling uncared for by therapist	1	0	1
2.3.2.2	Deference toward therapist	3	0	3*
2.3.2.3	Struggle to integrate therapy and life	3	0	3*
2.3.2.4	Lack of engagement	2	0	2
2.4	**Challenges in the process**	6	2	4
2.4.1	**Change of therapist**	6	2[a]	4
2.4.1.1	Non-interfering effect of change of therapist	5	2[a]	3
2.4.1.1.1	Positive perception	4	2[a*]	2
2.4.1.1.1.1	Like a new start	2	0	2
2.4.1.1.1.2	Handling it well	2	2[a*]	0
2.4.1.1.2	Beneficial impact	3	1	2
2.4.1.1.2.1	Second therapist better fit for client	3	1	2
2.4.1.1.3	No impact	3	2[a*]	1
2.4.1.1.3.1	Both / all therapists were experienced as helpful	2	1	1
2.4.1.1.3.2	Attached to process not therapist	1	1	0
2.4.1.2	Interfering effect of change of therapist	4	0	4*
2.4.1.2.1	Unfortunate timing for client	3	0	3*
2.4.1.2.1.1	In mid-process	2	0	2
2.4.1.2.1.2	Loss of support when feeling vulnerable	1	0	1
2.4.1.2.2	Disruption in building therapeutic relationship	2	0	2
2.4.1.2.3	Reinforcing expectation of being rejected in relationship	1	0	1
2.4.1.2.4	Difficulties in relationship with second therapist	1	0	1
2.4.2	**Delay / inconsistency in process**	4	1	3*
2.4.2.1	Inconsistency of sessions	3	1	2
2.4.2.1.1	Sense of abandonment	2	0	2
2.4.2.2	Took time to develop relationship	2	0	2
2.4.2.2.1	"Rocky start"	1	0	1
2.4.2.2.2	Took time to develop common understanding	1	0	1

[a] = maximum number of 'improvers' who had change of therapist.

3	**IN THE END**	N	↑	↓
3.1	**Ending of therapy**	8	4	4
3.1.1	**Facilitative aspects of ending**	4	4*	0
3.1.1.1	Feeling ready to end	4	4*	0
3.1.1.1.1	Able to continue progress made	2	2	0
3.1.1.2	Sessions helped a lot	4	4*	0
3.1.1.2.1	Made great progress	3	3*	0
3.1.1.2.2	Felt really good after therapy	2	2	0
3.1.1.3	Mutual decision to end	3	3*	0
3.1.2	**Interfering aspects of ending**	4	0	4*
3.1.2.1	Incomplete therapeutic process	4	0	4*
3.1.2.1.1	Client decision not to continue following therapist's decision to leave / take extended break	3	0	3*
3.1.2.1.2	Did not engage in ending process	2	0	2
3.1.2.1.3	Left feeling distressed and vulnerable	2	0	2

3.2	**Changes**	8	4	4
3.2.1	**Facilitative changes**	8	4	4
3.2.1.1	Improvement in key difficulties that wanted to address in therapy	8	4	4
3.2.1.1.1	Improvement during early stages in therapy	3	3*	0
3.2.1.2	Relationship with self	8	4	4
3.2.1.2.1	Increased self-awareness / self-understanding	7	4	4
3.2.1.2.1.1	More aware of needs & their impact	7	4	3
3.2.1.2.2	Increased self-acceptance	6	4*	2
3.2.1.2.2.1	More self-compassion	6	4*	2
3.2.1.2.2.2	Increased appreciation of self	3	3*	0
3.2.1.2.3	Feeling complete / more integrated / back on track	6	4*	2
3.2.1.2.3.1	Increased self-control / stillness / calm	4	3*	1
3.2.1.2.3.2	More access to emotions	2	1	1
3.2.1.2.3.3	More in touch with embodied self	1	1	0
3.2.1.3	Self in the world	7	4	4
3.2.1.3.1	More empowered / motivated	6	4*	2
3.2.1.3.1.1	Increased self-confidence / self-trust / self-belief	4	3*	1
3.2.1.3.2	New understanding of self in relationship	5	3	2
3.2.1.3.2.1	Less fearful of rejection and abandonment	4	3*	1
3.2.1.3.2.1.1	Able to deal with conflict in relationship	2	1	1
3.2.1.3.2.2	More able to put own needs first in relationships	3	2	1
3.2.1.3.2.2.1	Able to ask for needs to be met	1	1	0
3.2.1.3.2.2.2	Setting limits	1	1	0
3.2.1.3.2.2.3	Able to make choices for self	1	0	1
3.2.1.3.2.3	Increased awareness of impact on/of self in relationship	2	2	0
3.2.1.3.2.4	Figuring out who/how to trust	1	1	0
3.2.1.3.3	Increased openness	5	4*	1
3.2.1.3.3.1	To dealing with life as it comes	3	3*	0
3.2.1.3.3.1.1	Some things not in control	3	3*	0
3.2.1.3.3.2	To connecting with other people	3	2	1
3.2.1.3.3.2.1	More accepting of others' differences	2	1	1
3.2.1.3.3.3	To being in the moment	2	2	0
3.2.1.3.3.4	To taking risks / challenging self	1	1	0
3.2.1.3.4	More positive / balanced / realistic perspective on life	3	3*	0
3.2.1.4	Positive change in personal circumstances	4	3*	1
3.2.1.4.1	Increased resources	4	3*	1
3.2.1.4.1.1	Within relationships	3	2	1
3.2.1.4.1.2	Within self	1	1	0
3.2.2	**Interfering changes**	4	1	4*
3.2.2.1	Deterioration / no change in key difficulties	4	0	4*

3.2.2.1.1	Deterioration later in therapy	2	0	2
3.2.2.1.2	No clearer at end of therapy	2	0	2
3.2.2.1.3	Scores suggested did better when not in therapy	1	0	1
3.2.2.2	Disappointment with process	4	0	4*
3.2.2.2.1	Feeling worse	4	0	4*
3.2.2.2.1.1	More depressed / distressed	4	0	4*
3.2.2.2.1.1.1	Less accepting of self	2	0	2
3.2.2.2.1.1.2	Struggling to make peace with changes	1	0	1
3.2.2.2.1.2	More vulnerable	3	0	3*
3.2.2.2.1.2.1	Less willing to engage emotionally	2	0	2
3.2.2.2.1.2.2	Feeling less supported	1	0	1
3.2.2.2.2	Perception that getting nothing from the process	2	0	2
3.2.2.3	Loss of hope	4	0	4*
3.2.2.3.1	No resolution	3	0	3*
3.2.2.4	Negative change in personal circumstances	3	1	2
3.3	**Potential impact of research**	8	4	4
3.3.1	**Data may not be accurate**	7	4	4
3.3.1.1	May have wanted to please therapist / researcher	6	3	3
3.3.1.2	Hard to capture experience	5	3	2
3.3.1.2.1	Hard to shift between forms and session	2	1	1
3.3.1.2.2	Sometimes tired / unwell	2	1	1
3.3.1.2.3	No adjustments made	1	1	0
3.3.1.3	Self-conscious	3	1	2
3.3.1.4	Amount of paperwork hindering	1	0	1
3.3.2	**Increased self-worth**	3	2	1
3.3.2.1	By helping others	3	2	1
3.3.2.2	By helping self	3	2	1
3.3.3	**PQ provided motivation & focus**	2	1	1

Appendix K: Study 3 Analysis

Notes. ↑ = improvers; ↓ = deteriorators. *Italics* = meaning units. AB = Affirmative Brief; AR = Affirmative Rebuttal; AN = Affirmative Narrative; SB = Sceptic Brief; SR = Sceptic Rebuttal; SN = Sceptic Narrative.

1. In the beginning

1.1 Client concerns (8)

1.1.1 Self in the world (8)

1.1.1.1 Expressions of anxiety (7)

↑	Simon	*…anxiety […] recurring/ongoing […] causing him moderate difficulties. (SB)*
↓	Joseph	*Joseph reported difficulties with depression and <u>anxiety</u> as well as stress related symptoms which have persisted for > 12 months. (AN)*
↓	Sofia	*This is a theme which Sofia discusses throughout therapy, for example, the therapist notes in Session 6 indicate that the client added an item to her PQ "describing her anxiety about the pace of her life now in Glasgow in comparison to her experiences growing up in her home country". (AB)*

1.1.1.1.1 Stress/distress (2)

↑	Linda	*In Linda's first Change Interview, she states that her initial contact with the service was due to a temporary state of distress caused by a <u>stressful</u> work circumstance. (SB)*
↓	Joseph	*Joseph reported difficulties with depression and anxiety as well as <u>stress related symptoms</u> which have persisted for > 12 months. (AN)*

1.1.1.1.2 Social anxiety (2)

↑	Julia	*She reported experiencing long standing social anxiety difficulties which were hindering her daily life activities. (AN)*
↓	Luke	*His PQ items "I often feel anxious around other people"; "I'm inferior to other people"; "I feel quite isolated from other people"; and "I have no social life (almost)" all showed deterioration over the course of therapy. (SB)*

1.1.1.1.3 Existential anxiety (1)

↓	Sofia	*In Sofia's Intake Interview she identified the main problems that led her to seek therapy on the Problem Description Form . These included "existential anxiety around 'where is home?'" (AB)*

1.1.1.1.4 Panic (1)

↑	James	*"Yeah, I wasn't even aware actually that I was in a state of perpetual panic because I was so used to it." (AB)*

1.1.1.1.5 Psychosis (1)

↑	Simon	*… some of the issues that Simon presented with at the time of referral, such as medical diagnoses of depression and psychosis within the last 6 months, he rated as causing him moderate and mild difficulties respectively. (SB)*

1.1.1.2 Relationship with others (6)

↓	Joseph	*Joseph also reported recurring issues related to his self-esteem, relationships and work-related problems, all of which lasted > 12months in duration. (AN)*
↓	Joseph	*He described feeling stressed and concerned with how others perceived him and struggled to find a balance between work and free time. (AN)*
↓	Sofia	*she had a number of issues surrounding her transition to life in the UK, her relationship with her family and her self-confidence. (AN)*
↓	Luke	*When Luke began to email his therapist describing his home situation, it became clear that there were significant emotional difficulties in his home (SB)*
↓	Caitlin	*Difficulties in all relationships – family/friends/partner (Adjudication)*
↓	Caitlin	*She came to therapy having been depressed for a number of years and was experiencing difficulties in her relationships, including those with family, friends and her partner. When she came to therapy, she was struggling to cope with these difficulties but was hopeful that therapy could bring some form of resolution. (AN)*

1.1.1.2.1 Feelings of isolation (2)

↑	Linda	*PQ items: I feel I am invisible; I have no sense of belonging; I feel like an afterthought; I don't have anyone to fulfil my emotional needs; I don't feel appreciated; 10) I feel lonely (AB)*
↓	Luke	*PQ items: I often feel anxious around other people; I'm inferior to other people; I feel quite isolated from other people; I have no social life (almost). (SB)*

1.1.1.2.2 Tendency to take on other people's issues (1)

↑	Simon	*Session 19 HAT: Previously I had decided not to take on other people's issues so much because I found them overwhelming. (AB)*

1.1.1.3 Contextual difficulties

1.1.1.3.1 Work-related problems (2)

↑	Linda	*In Linda's first Change Interview, she states that her initial contact with the service was due to a temporary state of distress caused by a stressful work circumstance. (SB)*
↓	Joseph	*Joseph also reported recurring issues related to his self-esteem, relationships and work-related problems, all of which lasted > 12months in duration. Joseph described feeling 'held back' in his job and that his poor work-life balance was preventing him from doing the things he wanted. (AN)*

1.1.1.3.2 Adapting to life in UK (1)

↓	Sofia	*she had a number of issues surrounding her transition to life in the UK, her relationship with her family and her self-confidence. (AN)*
↓	Sofia	*difficulty adapting to cultural change (Adjudication)*
↓	Sofia	*In Sofia's Intake Interview she identified the main problems that led her to seek therapy on the Problem Description Form. These included "existential anxiety around 'where is home?'". (AB)*
↓	Sofia	*Her Personal Questionnaire also includes an item which states: "I feel inadequate in comparison with others in the UK" which she identified as being ongoing for one to two years. This is a theme which Sofia discusses throughout therapy, for example, the therapist notes in Session 6 indicate that the client added an item to her PQ "describing her anxiety about the pace of her life now in Glasgow in comparison to her experiences growing up in her home country". In addition, therapist notes*

		from Session 12 indicate that Sofia mentioned her "inner critic seemed to appear around the time she moved to the UK and has been getting louder ever since". (AB)

1.1.2 Relationship with self (6)

1.1.2.1 Depression (4)

↑	Simon	*…some of the issues that Simon presented with at the time of referral, such as medical diagnoses of depression and psychosis within the last 6 months, which he rated as causing him moderate and mild difficulties respectively. (SB)*
↑	Simon	*If life events or relationships get too much for Simon, his health and wellbeing can begin to suffer and so can his motivation for engaging in other activities such as those above. [For example, Session 10 HAT form] "I have had issues with funding and benefits recently and it has been stressing me out and tipping my balance into negative and my health was suffering, and I thought I might end up back in hospital" (SB)*
↓	Joseph	*Joseph reported difficulties with depression and anxiety as well as stress related symptoms which has persisted for > 12 months. (AN)*
↓	Luke	*Luke was in a very low place when he came to therapy, feeling very worthless and hopeless. (AN)*
↓	Caitlin	*She came to therapy having been depressed for a number of years (AN)*

1.1.2.2 Low self-esteem (4)

↑	Linda	*PQ items: II feel I am invisible; I feel like an afterthought; I don't feel appreciated. (AB)*
↓	Joseph	*Joseph also reported recurring issues related to his self-esteem, relationships and work-related problems, all of which lasted > 12months in duration. (AN)*
↓	Sofia	*she had a number of issues surrounding her transition to life in the UK, her relationship with her family and her self-confidence. (AN)*
↓	Sofia	*Her Personal Questionnaire also includes an item which states: "I feel inadequate in comparison with others in the UK" which she identified as being ongoing for one to two years. In addition, therapist notes from Session 12 indicate that Sofia mentioned her "inner critic seemed to appear around the time she moved to the UK and has been getting louder ever since". (AB)*

↓	Luke	*Luke was in a very low place when he came to therapy, feeling very worthless and hopeless. (AN)*

1.2 Readiness to begin therapy (8)

1.2.1 Motivated (7)

↑	Linda	*Change Interview: "But now I feel I want to bring them up, deal with them, and get rid of them. I do want that." (SB)*
↑	Simon	*it can be inferred that the motivation was an inherent aspect of his personality and it was not necessarily the therapy itself that brought it out (SB)*
↑	Simon	*Session 18 H.A.T form: "It's important to me that my family see I'm healthier + that I can affect their lives in positive ways". (AB)*
↑	Simon	*The above excerpt from the change interview suggests that Simon may be motivated to change for the benefit of his family (AB)*
↑	Simon	*A part of Simon's reason for being in therapy could be that he feels there is a need or pressure to recover from his mental health issues and get back to work. (SB)*
↑	Simon	*The cultural and/or the personal expectation Simon placed on himself could have been a part of his "must try harder" or have more motivation and in interpreting or perceiving his experience of being in therapy to be a part of what it means to try harder. In terms of the sceptic brief even if Simon had developed these of conditions of worth in relation to work and benefits, they could still be a part of him which was motivating him to change. (SB)*

1.2.1.1 Determination (5)

↑	Linda	*Throughout Linda's Change Interviews,) she repeatedly indicated an expectation and a personal determination to change. (SB)*
↑	Linda	*Additionally, this determination could be viewed in relation to Linda's extensive self-help and extra-therapeutic efforts made throughout therapy that could hold causal significance in her change (SB)*
↑	Linda	*Linda was determined to come to therapy to work on her problems – she took a determined and proactive approach (Adjudication)*
↑	Simon	*…was a very determined individual (Adjudication)*

↑	Simon	*"I think, determination to keep going." (SB)*
↓	Luke	*The client appeared to be determined to continue with the therapeutic process, showing agency in seeking appointments, and communicating extensively over email (Adjudication)*

1.2.1.1.1 Stubborn (2)

↑	Julia	*desire to make a shift, indeed she identifies her own stubbornness and how this may influence her desire to change (Adjudication)*
↓	Caitlin	*Stubbornness (Adjudication)*

1.2.1.2 Expectations (4)

↑	Linda	*Having accurate and optimistic expectations would invariably enable Linda to utilise the therapy to its fullest. (AR)*
↑	Linda	*Change Interview: "Yeah, because I knew I wasn't dealing with things, and I think that was a part of coming here... Knowing that it will hopefully give me the tools that will enable me to deal with things properly and process them." (p.15) (SB)*
↑	Linda	*Throughout Linda's Change Interviews, she repeatedly indicated an expectation and a personal determination to change. This shows that she had significant expectations from the therapeutic process, and a strong sense of personal agency in making changes. (SB)*
↑	Simon	*Session 27 HAT: "So I can develop more coping mechanisms"(AB)*
↓	Joseph	*Change Interview: "maybe I had the wrong expectations" (SB)*
↓	Caitlin	*Furthermore, by being a participant and 'giving something back', we can infer from this that there may have been an expectation to gain something for herself through therapy – such as a resolution for the difficulties she was experiencing at the time, which has been previously discussed earlier in this rebuttal. (AR)*
↓	Caitlin	*Another notable hindering aspect shared in this interview was that she had been; "…unpacking stuff but no resolution yet."* *This is an indication of Caitlin's expectation of the therapeutic outcome based upon how much of herself she had shared in therapy. (AB)*
↓	Caitlin	*Furthermore, she noted a change in her first Change Interview to be 'Not miserable anymore' and rated it as 'Expected', this*

		points towards an expectation of hers that therapy would continue to alleviate the misery that she had felt. (AR)

1.2.1.3 Hope (3)

↑	James	*James worked hard and invested himself in therapy stating that "he made the decision to trust his therapist," even though trusting people was difficult for him. (AB)*
↑	Simon	*Session 1 HAT: "I noticed that I am hopeful for this process and that my PQ scores are lower than they were at my initial interview" (AB)*
↓	Caitlin	*When she came to therapy, she was struggling to cope with these difficulties but was hopeful that therapy could bring some form of resolution. (AN)*

1.2.2 Open to experience (6)

1.2.2.1 Able to be reflective (5)

↑	Linda	*She was reflective and processed lots outside and inside therapy (Adjudication)*
↑	Linda	*tendency to self-process outside of the session (AR)*
↑	James	*he was able to articulate his thoughts clearly and explain what his processes were (Adjudication)*
↑	Simon	*"There's also other things to do with relationships where I know personally where I'm quite reflective, empathic and assertive about the different things." (SB)*
↑	Simon	*"I think that I've got a certain way of thinking, whether it's the career or it's a natural reflection. What started it was an incredible shyness, introspection and then someone saying I was reserved." (SB)*
↓	Sofia	*she was able to be self-reflective (Adjudication)*
↓	Caitlin	*willing to use sessions to be open and reflective. (AB)*

1.2.2.2 Willing to be open (3)

↑	Julia	*an openness and acceptance of therapy (Adjudication)*
↑	Linda	*Linda was willing to engage with therapy (Adjudication)*
↑	Linda	*readiness to explore specific issues in therapy (AR)*
↓	Caitlin	*Apparent willingness to engage in the therapy – as per emails/HAT processes/even the sad ending of her feeling vulnerable and opened-up (Adjudication)*

↓	Caitlin	*The extensive content of [her HAT forms] as well as other qualitative data obtained from Change Interviews support the notion that Caitlin was invested in therapy and willing to use sessions to be open and reflective. (AB)*

1.2.2.3 Prepared to go to 'unhappy places' (2)

↑	Linda	*Change Interview: "No, I think you've got to go there... I think I was aware that with therapy, you have to go to the unhappy places to get resources." (SB)*
↓	Caitlin	*Caitlin had an expectation of increased feelings of vulnerability because of attending counselling (AB)*

1.2.3 Prepared (4)

1.2.3.1 Ready to get started (2)

↑	Julia	*desire to make a shift (Adjudication)*
↑	Linda	*Change Interview: "But now I feel I want to bring them up, deal with them, and get rid of them. I do want that." (SB)*
↑	Linda	*Change Interview: "Yeah. I think I came from the point of view that it's like... Well, I'm ready to make the changes. I'm ready to change for me." (SB)*
↑	Linda	*Change Interview: "So, it was like erm... No definitely, I was ready for change when I came here. I knew I had to change. If I was to go on, then I had something... Something had to change." (SB)*

1.2.3.2 Proactive (2)

↑	Linda	*Linda was determined to come to therapy to work on her problems – she took a determined and <u>proactive</u> approach (Adjudication)*
↑	Simon	*Session 15 HAT: "It's good to be back! I can't remember when I was last here but feel positive to meet new therapist. It was combination of things, but I think it was good of therapist to remind me that it's up to me to choose what I want to talk about in the sessions. <u>I'm going to take copies of PQ's away with me and make some notes for next session.</u> Today was good to be back and reminds me of the process that takes place here" (AB)*

1.2.3.2.1 Using other therapeutic services (2)

↑	Linda	*Some of these include Linda's use of other therapeutic services, such as holistic and spiritual healers. These appear to fall within pivotal sessions such as Linda's 'Inner Child' session; potentially causing or contributing to the facilitation of this important in-therapy event where Linda's PQ score fell to indicate no distress at all. Linda's therapist also notes the personal value she placed in these services (SB)*
↑	Simon	*Session 6 HAT: " I had just come from music therapy and it had helped me feel much better. It made me realise that the therapeutic classes I have can really change the way I feel and that I am getting progressively healthier." (AB)*

1.2.3.3 Knew what wanted to work on (2)

↑	Linda	*She could identify her own needs (Adjudication)*
↑	Linda	*"And then there came a point when I thought, no, maybe there are issues that I really need to bring up, release, and let go in this safe environment. And really deal with things… (AB)*
↑	Linda	*HAT Session 3: "I spoke of an incident, that I have kept locked away. I need to deal with the issue, as it holds me back from being happy." (AB)*
↑	Linda	*HAT session 5. "My conversation relating to my marriage break up. I didn't really know what happened or what I did to make my marriage break up. I need to understand." (AB)*
↑	Linda	*her preparation and readiness to explore specific issues in therapy (AR)*
↓	Sofia	*knew what she wanted to work on (Adjudication)*

1.3 Potential challenges to readiness (6)

1.3.1 Doubt / potential for doubt (4)

↑	Linda	*Change Interview: "I think so... I think when I came here, I didn't need it […] Oh no, I definitely felt like I didn't need it... I didn't feel like I needed it at all […] So, I thought it was really just the job. It really was just the pressure. I didn't like where I was, so I felt much better. So, by the time I came to coming here, I wasn't going to come, because I thought I didn't need it now." (SB)*
↓	Luke	*He was doubtful that therapy could help him (AN)*

1.3.1.1 Expectation that would feel more vulnerable) (1)

↓	Caitlin	*Caitlin had an expectation of increased feelings of vulnerability because of attending counselling (AB)*

1.3.1.2 ‘Wrong’ Expectations (1)

↓	Joseph	*CI: “maybe I had the wrong expectations but it was, from the beginning pretty clear that it’s me doing all the talking. Sometimes it feels like, you kinda want that advice or that constructive feedback, a bit more honest” (SB)*

1.3.2 Difficulties in processing emotions and experience (4)

1.3.2.1 Difficulty in accessing and expressing emotions / experience (3)

↑	James	*The boulder that was shutting my emotions - I had to get rid of the boulder (P98) (AB)*
↑	James	*James gained more access to, and awareness of his emotions: “I’ve been able to open up my emotional world – even though emotions are awkward.” (AB)*
↓	Joseph	*He continued to find it difficult to symbolise his experiences. (SB)*
↓	Joseph	*thus being unable to find ways of expressing his process. (SB)*
↓	Luke	*and it seems very plausible that Luke had very little emotional maturity or awareness. (SB)*
↓	Luke	*Luke was extremely detached from his own experience, writing “When I try to talk about such things it feels to me as if I was somebody else’s story.”. This hindered him from bringing content into therapy; he wrote that he could “barely remember what exactly happens and it’s even harder to explain to somebody else.” (SR)*
↓	Luke	*Therapist notes: "no thoughts, no feelings";"doesn't seem to have feelings";"so out of touch with himself";"it seems that accessing his emotions is quite threatening to him and he 'blanks' out all awareness of them" (SB)*
↓	Luke	*He therefore would have struggled to articulate and understand how he was benefitting from therapy; perhaps latterly all he was able to understand was a growing anger and dissatisfaction. (SB)*
↓	Luke	*and it seems very plausible that Luke had very little emotional maturity or awareness. (SB)*

1.3.2.2 Difficulty understanding own experience (2)

↑	James	*Change Interview: “I used to think most of my time in some part of my head which was connected to outer space.” (AB)*
↓	Luke	*He therefore would have struggled to articulate and understand how he was benefitting from therapy; perhaps*

		latterly all he was able to understand was a growing anger and dissatisfaction. (SB)

1.3.2.3 Wanting to avoid difficult emotions (2)

↑	Simon	*Session 19 HAT: "I don't really want to talk too much about that because it tips my balance to painful negativity just talked briefly about it. What made this important was I didn't allow it to take over + and keep hurting me and focused on more positive events that had more positive value for me and the people I love. Trying to keep life simple and find support for difficult areas." (AB)*
↑	Simon	*We felt Simon had a strong desire to focus on the positives which may have prevented him from looking too clearly at the negatives (Adjudication)*
↑	Simon	*Session 12 HAT: "<u>I wanted to keep things positive</u> so, I talked about my birthday last week and showed the therapist pictures of my birthday weekend and explained how much it meant to me." (AB)*
↓	Luke	*Therapist notes: "he has such a strong resistance to acknowledging any difficult emotion" (SB)*

1.3.3 Tendency toward rigidity (2)

1.3.3.1 Perfectionism (1)

↑	James	*The client's perfectionism and ability to think logically and clearly helped him to process week to week (Adjudication)*

1.3.3.2 Unwillingness to consider/accept another's point of view (1)

↓	Caitlin	*The most frequently referred to hindering personal characteristic expressed by Caitlin was that she <u>believes herself to be stubborn and unwilling to accept or consider another individual's point of view</u>, this was stated in her 1st, 2nd and 3rd Change Interviews. (SB)*

1.3.4 Lack of supportive relationships (2)

↓	Luke	*Luke's circumstances were hindering as well, notably his feelings of isolation. This was demonstrated by him spending his 20th birthday alone, and a growing awareness that he was going to fail his exams because he had not been attending lectures and had nobody to ask for notes. (SR)*
↓	Luke	*Furthermore, Luke's own self-reporting that he was not engaging in his classes, shows that this disengagement was not unique to therapy and was an aspect of Luke's approach to*

		everything at that time. He was also distant from his peers and from his family. (SB)
↓	Caitlin	*As well as being frequently explored during therapy, her relationships with others were also mentioned in Change Interviews as limiting factors to Caitlin's experience of therapy. (SB)*

2. In the process

2.1 Therapeutic climate (8)

2.1.1 Safe (6)

↑	Linda	*Linda provides extensive qualitative description of how exploring this pervasive issue in the safe environment with her counsellor enabled her to open up and work through this issue outside of therapy (AB)*
↑	Linda	*Change Interview: "And then there came a point when I thought, no, maybe there is issues that I really need to bring up, release, and let go in this safe environment. And really deal with things…" (AB)*
↑	James	*HAT 19: "I felt that it was safe to share/express my feelings in a more emotional way and not just in an academic way." (AB)*
↓	Luke	*The second therapist offered him a safe space to be angry, allowing him to release what had been bottled up inside him; what had been destroying him. She accepted him back, week after week – not making herself an enemy no matter what Luke did. He expected people to turn on him eventually, stating "It seems that every person I've ever met became my enemy". It was probably incredibly healing for Luke to have someone accept him despite his anger towards them. (SB)*
↓	Caitlin	*In support of the above evidence, Caitlin's experience of helpful therapy processes addressed in Change Interviews gives us a strong indication that Caitlin found her therapists to be 'empathic', 'understanding' and 'safe'. (SB)*
↓	Caitlin	*Helpful: Support that she received – feeling safe at the research clinic (Adjudication)*
↓	Caitlin	*A very powerful statement made by Caitlin in her HAT form from her final session which is an appropriate final supporting point for this particular argument is: "It was a safe space so I could be me, exactly me, and not an act." (SB)*
↓	Caitlin	*She explicitly highlights aspects of therapy which have been beneficial to her, such as the importance of the space being 'safe' to 'project thoughts' and 'open up'. (SB)*

2.1.1.1 Able to express feelings without upsetting someone else (2)

↑	Linda	*HAT Session 4: "I was able to express my feelings without upsetting someone else." (AB)*
↓	Joseph	*When asked to identify helpful aspects in the therapeutic process Joseph is able to offer several examples which relate predominantly to the opportunity for him to talk and share his experiences with a neutral person. (SB)*

2.1.1.2 Able to be real / authentic (model for how relationships can be) (2)

↑	Simon	*The therapeutic relationship was real and helped Simon to feel that he and others close to him could work through relational difficulties. "It's the thing about having an authentic relationship." (AB)*
↓	Caitlin	*The qualitative data extracted from Caitlin's HAT forms and Change Interviews provide concrete evidence that Caitlin experienced two therapy relationships in which she felt comfortable to be authentic and understood (SR)*
↓	Caitlin	*Helpful: Felt able to be herself in the process (authentically) (Adjudication)*
↓	Caitlin	*A very powerful statement made by Caitlin in her HAT form from her final session which is an appropriate final supporting point for this particular argument is: "It was a safe space so I could be me, exactly me, and not an act." (SB)*

2.1.1.3 Able to experience relational/emotional depth (1)

↑	Linda	*A specific therapy process that Linda identified as being extremely helpful was the exploration or 'letting in' of her inner child, which she described as making her feel complete. This process was identified as a profound moment of relational depth between Linda and her therapist, where every item on the Relational Depth Inventory (Wiggins, 2013), apart from 'I felt as if time had stopped' which she described as 'Very Much', was rated as 'Completely' (AB)*
↑	Linda	*Therapist notes: "We agreed to end after our next session, and both acknowledged that something huge had happened that couldn't be explained". (AB)*

2.1.1.4 Able to disclose / show vulnerability (2)

↑	Linda	*Showing vulnerability. (AB)*
↓	Caitlin	*In further support of the above evidence, in her HAT forms Caitlin also revealed something to her therapist which was*

		important to her on 31 of her sessions; furthermore the majority of these disclosures were rated as either 'greatly' or 'extremely' important. The frequency with which Caitlin felt safe and comfortable enough to disclose important content suggests good quality and trusting therapist-client relationships. (SB)

2.1.2 Warmth / connection (6)

↑	Julia	*HAT session 1: "My counsellor has a really good way of putting words to it." (AB)*
↑	James	*HAT session 14: "Since the last session I had become much more aware of my therapist as a person." (AB)*
↑	Simon	*Session 10 therapist notes: "He said he wanted me to see it as I had mentioned before that there were parts of his life I didn't know about it. I was quite touched that he made such a big effort." (AR)*
↑	Simon	*Session 32 therapist notes: "Final session with writer. Client reports feeling very tired today. Had given thought about how he wanted to 'end' and shared a clip from YouTube and a text message which had strong personal meaning for him." (AR)*
↑	Simon	*Session 33 therapist notes: "Felt very much like we were creating a foundation together." (AR)*
↑	Simon	*Session 37 therapist notes:" I felt connected to him and less like I needed to prove my understanding to him." (AR)*
↓	Sofia	*Helpful: the love and care of her first therapist (Adjudication)*
↓	Luke	*Luke's last emails to and from the second therapist, as well as the notes from the last session, demonstrate a degree of warmth and prizing. (SR)*
↓	Luke	*…they ended therapy on a very warm note, laughing together and showing regret that they didn't end properly. Therefore, Luke's behaviour following the ruptures is much more consistent with the theory that Luke was, on one level, finding therapy valuable, he prized and trusted his therapist, and the ruptures were a sign of his progress. (SR)*
↓	Caitlin	*Session 47: Responds positively to counsellor's empathy regarding her anger.* *Session 48: Responds positively again to counsellor's empathy. (SB)*

2.1.3 Supportive (6)

↑	Julia	*therapy seems to be a supportive environment for Julia (Adjudication)*

2.1.3.1 Space to process (2)

↑	Simon	*The biggest help that we could see was the therapeutic relationship and the offering of a space where Simon could process what was important to him (Adjudication)*
↑	Simon	*The helpful events that seem to be most common in terms of Simon's experience is having the space and freedom to discuss and focus on what seems to be important for him in relation to where and how he is personally as well as what is going on his life (AB)*
↓	Joseph	*When asked to identify helpful aspects in the therapeutic process Joseph is able to offer several examples which relate predominantly to the opportunity for him to talk and share his experiences with a neutral person, that he has space and time to work through things that are in his head, that he is challenged to think differently about aspects of his process and that he feels his experience is validated. (SB)*

2.1.3.2 Opportunity to practice communication skills (2)

↑	James	*James and his therapist practiced communication skills (AN)*
↑	Simon	*Session 43 therapist notes: "Anticipation/preparation of his talk with daughter. After session showed me objects/ photos/ films/ books that he is going to use. Spoke as if he was speaking to her. Named feelings, hopes, needs, & fears" (AR)*

2.1.3.3 Support between sessions (2)

↓	Sofia	*Additionally, Sofia has contact with the first therapist outwith the session; the therapist sends her articles and links to material she thinks may be relevant to Sofia and Sofia also contacts the therapist regarding a medical condition she is concerned about. This suggests that the strong connection she has with her therapist extends outwith the 50 minute session. (AB)*
↓	Caitlin	*Additionally, to state that she was not accommodated for is not strictly true. Caitlin's therapist emails back promptly which to offer a session via telephone which unfortunately Caitlin does not feel able to accept; Caitlin does however respond positively to the immediacy of her therapist's response, which was a comfort to her. (SR)*

2.1.4 Collaborative (4)

↑	Julia	*She mentioned "we" on a couple of occasions which indicates the collaboration between client and therapist (Adjudication)*
↑	Simon	*Session 8 therapist notes: "He mentioned that he had noted down comments that I had made previously about perhaps getting into more depth in our sessions and slowing things down a little.* *I said this was interesting as I wanted to have a review of our sessions to see whether he was getting what he wanted from them.* *I told him that from my perspective it felt a bit like that he was staying on the surface and talking about many things, but not in any depth. I said it was sometimes difficult to get in to say anything as he was talking quickly. I said that this was okay if it was what he wanted and if it was benefitting him.* *He said he had found our sessions helpful but could understand what I was saying. He said that he would be up for trying to get more out of the sessions. I said that in future I might try and come in more often if I felt there were issues he was touching on that I felt would benefit from being developed further. He said that was fine.* *I felt the rest of the session was better as we touched on things that we hadn't discussed before." (AR)*
↑	Simon	*Session 10 therapist notes: "He said he wanted me to see it as I had mentioned before that there were parts of his life I didn't know about it. I was quite touched that he made such a big effort." (AR)*
↑	Simon	*Session 32 therapist notes: "Had given thought about how he wanted to 'end' and shared a clip from YouTube and a text message which had strong personal meaning for him." (AR)*
↑	Simon	*Session 33 therapist notes: "Felt very much like we were creating a foundation together." (AR)*
↓	Joseph	*Change Interview: "…but that was good because it made me go; what do you want me to talk about, what do you want me to say?" (AR)*
↓	Joseph	*Change Interview: "…the first session she was like 'oh I feel like we go round in circles' and I went 'oh do you mean we go round in tangents', she just said she wanted me more to focus on something instead of jumping from incident to problem to you know feelings, which has been good. I mean, I feel like, with therapist 2, I feel like there is more of a tactic going on,*

		there's more of a kind of process going on, like I don't feel like I just came in here to talk and then walk away."
↓	Caitlin	*Caitlin expresses in her HAT (Appendix A Table A3) forms various therapy processes which directly involve the therapist. (SB)*

2.2 Therapist in the process (8)

2.2.1 Facilitative therapist processes

2.2.1.1 Accepting/understanding/validating client (8)

↑	James	*The non-judgemental presence of the therapist seemed to be helpful to the client (Adjudication)*
↑	James	*…the acceptance of the therapeutic relationship (Adjudication)*
↑	James	*James explicitly expressed his difficulty understanding and feeling understood by others. Therapy gave James a forum to feel accepted and understood. (AB)*
↑	James	*Therapist acceptance allowed James to reduce feelings of guilt and shame and he was able to have a rewarding interpersonal experience. (AN)*
↓	Sofia	*Helpful: being able to talk, <u>be accepted</u> and listened to (Adjudication)*
↓	Sofia	*…no indication that Sofia was uncomfortable with the new therapist. The qualitative data from the HAT forms in fact suggests that the new therapist was able to hold, accept and empathise with Sofia's experience (SR)*
↓	Luke	*Helpful: The second therapist seems to regard the client unconditionally, despite the apparent anger expressed in sessions (Adjudication)*
↓	Luke	*She accepted him back, week after week – not making herself an enemy no matter what Luke did. He expected people to turn on him eventually, stating [in an email] "It seems that every person I've ever met became my enemy". It was probably incredibly healing for Luke to have someone accept him despite his anger towards them (SB)*
↓	Luke	*Luke had been difficult to engage throughout therapy, providing little content or context. Towards the end of therapy, however, he began to engage, in his own terms, in a way that he could. His actualising tendency at work (Rogers, 1951), he was enabled to express himself to <u>someone who accepted him</u> (SB)*
↓	Luke	*Helpful: The second therapist seems to recognise and appreciate the client's process. For example, acknowledging*

		that every expression of seeing the client is a threat to him (Adjudication)
↓	Caitlin	*…further correlates with what Caitlin expressed at her Mid-4 Change Interview, that her therapist is "empathic and understanding" (SB)*
↓	Caitlin	*While not explicitly referred to in the affirmative argument, it is important for the process of making the sceptic case as robust as possible that even any vague suggestion of therapist pressure must be dispelled. There is no evidence at all to suggest that pressure to 'open up' experienced by Caitlin was placed upon her by either of her therapists from any of the available data and the lack of session recordings means that this suggestion cannot be confirmed. (SR)*

2.2.1.1.1 Working within client's frame of reference (5)

↑	Linda	*Therapeutic relationship – therapist worked within Linda's frame of reference, allowing for exploration (Adjudication)*
↑	Linda	*In Linda's therapist's 'Unusual Within-Therapy Events', she claimed: "Session 18: I reflected that there was one tree (according to client's image) that still needs to be knocked down, and added that I don't want to push the client to do that and hope I haven't today." (AB)*
↑	Simon	*Session 15 HAT: "I think it was good of therapist to remind me that it's up to me to choose what I want to talk about in the sessions." (AB)*
↑	Simon	*Simon very rarely specifically spoke about a specific instance in the therapy or about something the therapist's said or did that was helpful. It was for the most part that Simon was able to use the therapy time and space as he wanted. (AR)*
↑	Simon	*Session 7 Therapist notes: "Simon was feeling tired today and was not as talkative as normal. He did his usual thing of talking about various subjects and not leaving much space for me to get in. However, after about 40 mins he went quiet. I let him stay quiet and then we talked a little bit about why he was quiet." (AR)*
↓	Sofia	*In session 19, she also wrote about how comfortable she felt suggesting things in therapy: "the fact that I suggested a new way of dealing with my concerns - comfortable enough to feel that this is something I can place as a proposal in the therapy space [sic]". (SB)*

2.2.1.1.1.1 Using metaphor (2)

↑	Julia	*HAT session 4: "The part about navigating anxiety being like an acrobat on a bicycle on a wavering line, where falling would be falling into anxiety" (AB)*
↑	Julia	*HAT session 1: "Talking about it makes me realize some feelings I couldn't label before. My counsellor has a really good way of putting words to it" (AB)*
↑	Linda	*In Linda's therapist's 'Unusual Within-Therapy Events', she claimed: "Session 18: I reflected that there was one tree (according to client's image) that still needs to be knocked down, and added that I don't want to push the client to do that and hope I haven't today." (AB)*

2.2.1.1.1.2 Reflecting words back (1)

↑	Simon	*HAT session 18: "Maybe about half way through the therapist put my words back to me, that I was saying I was incredibly overwhelmed by health issues because they were so heavy and all-consuming and now I want to feel more rational." (AB)*
↑	Simon	*Session 34 therapist notes: "Looking to the future, speaking motivationally (I repeated this to him-we discussed this)." (AR)*

2.2.1.1.1.3 Reflecting feeling (1)

↓	Caitlin	*Session 11: Counsellor reflects how much anger is present for Caitlin. (SB)*

2.2.1.1.2 Validating/affirming client's experience (4)

↑	Simon	*Session 15 HAT: " I talked freely for some time and therapist agreed that there was a lot going on. This helped me to decide to look at my life and decide that some things are in the past." (AB)*
↓	Joseph	*When asked to identify helpful aspects in the therapeutic process, Joseph is able to offer several examples which relate predominantly to the opportunity for him to talk and share his experiences with a neutral person, that he has space and time to work through things that are in his head, that he is challenged to think differently about aspects of his process and that <u>he feels his experience is validated</u>. (SB)*
↓	Joseph	*Change Interview: "but it wasn't my intention to upset her but it validated that what i was saying out loud, that it was mattered and it was upsetting" (SB)*
↓	Sofia	*Helpful: being able to talk, be accepted and listened to (Adjudication)*

↓	Sofia	*With the support of the therapy, she was able to act out of her personal agency and was affirmed in validating and experiencing her ability to make choices for herself (Adjudication)*
↓	Caitlin	*The qualitative data extracted from Caitlin's HAT forms and Change Interviews provide concrete evidence that Caitlin experienced two therapy relationships in which she felt comfortable to be authentic and understood (SR)*

2.2.1.1.3 Helping client feel comfortable (3)

↑	Simon	*Session 1 HAT: "It was important that the therapist made me feel comfortable with being involved in the process." (AB)*
↓	Sofia	*…no indication that Sofia was uncomfortable with the new therapist. The qualitative data from the HAT forms in fact suggests that the new therapist was able to hold, accept and empathise with Sofia's experience (SR)*
↓	Sofia	*It is evident through the aspects Sofia identified as helpful throughout therapy that she is comfortable with her therapist. Indeed, she wrote in the Helpful Aspects of Therapy form after session 16 that it was helpful for her therapist to be friendly and supportive as it allowed her to: "rest on something positive since my own experience has been so destructive" (SB)*
↓	Sofia	*In session 19, she also wrote about how comfortable she felt suggesting things in therapy: "the fact that I suggested a new way of dealing with my concerns - comfortable enough to feel that this is something I can place as a proposal in the therapy space [sic]". (SB)*
↓	Caitlin	*The qualitative data extracted from Caitlin's HAT forms and Change Interviews provide concrete evidence that Caitlin experienced two therapy relationships in which she felt comfortable to be authentic and understood (SR)*

2.2.1.2 Offering opportunities (5)

2.2.1.2.1 To ground / slow down client (3)

↑	Julia	*HAT session 4: "Also the feeling of being present in my body and mind at the same time as having anxiety or feeling panic" (AB)*
↑	Julia	*HAT session 8: "Talking about getting off my anti-depressants had me feeling really anxious about it in the session, but then we turned it around and I was reminded that I do know how to ground myself in an anxious situation." (AB)*
↑	Simon	*Session 7 Therapist notes: "I said that if he wanted we could slow things down a bit. I said it was really up to him how the sessions were inducted as he was in control. He commented*

		about being here before and that seemed to be prompting some thinking for him." (AR)
↓	Joseph	*Change Interview: "..yeah, kind of anchors me back to what she wants to talk about or what she wants to know more about, that kind of thing." (AR)*

2.2.1.2.2 To develop process (3)

↑	Linda	*Linda's post-therapy Change Interview provides a qualitative account of this profoundly important therapeutic process, attributing this phenomenon to the counsellor offering the opportunity to allow the inner child back in. (AB)*
↑	James	*HAT 34: "My therapist was able to make many connections today between the different threads of our discussion." (AB)*
↓	Sofia	*Additionally, on the HAT forms from sessions 12 through 15 Sofia mentions suggestions made by her first therapist which she rates as extremely helpful. (SB)*

2.2.1.2.3 To engage client (2)

↑	Simon	*Change Interview: "I was coming here just, probably offloading a hell of a lot of negative stuff and she [the therapist] just chipped in at one point, 'I've never heard you talk about your dad before', and at this point it was really pivotal. I just went, 'wow', and it hit me emotionally, I says 'WOWW! I've been coming in here and just moaning and moaning, and then started to change the thing of, but something definitely shifted [in my way] of looking at my life." (AB)*
↑	Simon	*Session 9 therapist notes: "I still find that he wants to go into a lot of detail about what is going on a daily basis and interrupt him a bit more to try and focus on areas that are a bit deeper." (AR)*
↑	Simon	*Session 42 therapist notes: "Found more gentle ways of non-verbally coming into check if it's okay for me to say something-not wanting to interfere/stop his process but wanting to offer something." (AR)*
↑	Simon	*Session 34 therapist notes: "Looking to the future, speaking motivationally (I repeated this to him - we discussed this)." (AR)*
↓	Luke	*Both therapists tried numerous times to engage Luke in therapy and to explain how the process worked, and Luke himself acknowledged to the first therapist in an email that he knew he hadn't been engaging. (SB)*

↓	Luke	*Helpful: The second therapist seemed to something creative and different that enabled the client to express his feelings and agency (Adjudication)*

2.2.1.2.4 To challenge client (2)

↑	Julia	*HAT session 2: "This idea that I don't want to accept my anxiety as part of me got challenged. Maybe I have to accept it in order to get past it? It's something I'm going to think about" (AB)*
↓	Joseph	*Helpful: the therapist challenging client in the therapy session (Adjudication)*
↓	Joseph	*Change Interview: "So, it's not that she's giving me advice but she makes me challenge myself a bit more, because sometimes I feel like, with therapist 1, I was coasting a wee bit, it was like I was turning up and just talking and going away and not thinking about it, whereas with therapist 2 I left having to think about it a bit more" (SB)*
↓	Joseph	*Change Interview: "I think therapist 2 is good because she goes after things and challenges things." (AB)*

2.2.1.3 Being transparent (5)

2.2.1.3.1 Responding to client's questions (3)

↑	Simon	*Session 7 Therapist notes: "He also asked at one point in the session how I thought he was doing in the sessions. I said that I found he just liked to talk about many things usually related to the week he has had and about all the classes he attends." (AR)*
↑	Simon	*Session 7 HAT: "I asked the therapist how I was doing. This was helpful as my health problems cause problems with my energy and sometimes I'm not so positive - it helped to hear what the therapist had to say." (AB)*
↑	Simon	*Session 16 HAT: "The new therapist answered my questions about a number of different things to do with counselling that I'd forgotten because I was away for a while. I feel more positive and looking forward now." (AB)*
↑	Simon	*Change Interview: "One thing that [third therapist] said that I was definitely aware that you wouldn't give advice because that's another thing. So, I had found a way to word things towards her and eh there was a few times where she ended up writing out some quotes what counselling or therapy meant to her, so I could use this for my daughter." (SB)*

↓	Joseph	*Change Interview: "…but that was good because it made me go; what do you want me to talk about, what do you want me to say?" (AR)*
↓	Sofia	*Furthermore the emails from Sofia suggest that she has developed a strong dependence on her therapist's opinion which may mean she would be less likely to rate her low on the TRS: "I wanted to get your advice on this; is this something that we could look at together and discuss during our next appointment or should I consult my GP and get some medical treatment for it? … I just wanted to get your advice/suggestion on the matter first. Of course, I will be happy to talk you through it (habitual pattern, duration, symptoms etc [sic]) when I see you next." (AR)*

2.2.1.3.2 Discussing process (4)

↑	Linda	*In Linda's therapist's 'Unusual Within-Therapy Events', she claimed: "Session 18: I reflected that there was one tree (according to client's image) that still needs to be knocked down, and added that I don't want to push the client to do that and hope I haven't today." (AB)*
↑	Simon	*Session 7 Therapist notes: "I asked him how he thought things were going and he said that he benefitted from the sessions as there was nowhere else he could process everything that had happened in the week. More discussion about what he wants from sessions." (AR)*
↑	Simon	*Session 8 therapist notes: "He mentioned that he had noted down comments that I had made previously about perhaps getting into more depth in our sessions and slowing things down a little. I said this was interesting as I wanted to have a review of our sessions to see whether he was getting what he wanted from them." (AR)*
↑	Simon	*Session 27 HAT: "I discussed what I was getting out of sessions and what works with therapist + me." (AB)*
↑	Simon	*Session 33 therapist notes: "Our first session. Established basics together. Practicalities. I acknowledged that he has experienced a few different counsellors and that we only have 7 sessions together: decided we would meet fortnights to cover longer as he is busy and suffers from lack of energy." (AR)*
↑	Simon	*Session 33 Therapist notes: "I asked what he would like from this time. He talked through his PQ items, spoke about how things have shifted for him. I spoke about what he is looking towards, what he would like, what's important to him, what he wants. More to the future." (AR)*

↑	Simon	*Session 39 therapist notes: "Talk about what happens when we go onto something else-what he needs?"*
↓	Joseph	*Change Interview: "…but that was good because it made me go; what do you want me to talk about, what do you want me to say?" (AR)*
↓	Joseph	*Change Interview: "…the first session she was like oh I feel like we go round in circles and I went oh do you mean we go round in tangents, she just said she wanted me more to focus on something instead of jumping from incident to problem to you know feelings, which has been good, I mean, I feel like, with therapist 2 I feel like there is more of a tactic going on, there's more of a kind of process going on, like I don't feel like I just came in here to talk and then walk away." (AR)*
↓	Luke	*Early on, his therapist noted that Luke was getting something from the process, noting in session 5 "he seems to be getting something from this" and in session 7 "merely spending this one hour per week reflecting or experiencing is good for him". In session 14 the notes report that Luke "responded negatively to the suggestion that he feels obliged to attend, and unequivocally to the question that he is getting something from the process". (SB)*

2.2.1.3.3 Noticing / validating changes (2)

↑	James	*HAT 25: "My therapist noted some things where I had been assertive and also pointed out that I could feel quite proud of how I had handled myself with my mother." (AB)*
↑	Simon	*Session 3 HAT: "I was having a bad day and the therapist said well done for getting here and for using your bike for the first time this year. This was important because it came from the therapist." (AB)*
↑	Simon	*Session 4 HAT: "The therapist told me I looked well and that she could see a change in me." (AB)*
↑	Simon	*Session 35 HAT: "I came in tired […] + <u>at end therapist said, "you're holding it together". That made a big difference to me</u> as the last two weeks have been quite a lot to get through but have made it + this session gives me a boost of positivity + time to realise my vision as I write this and know I have plans to do good things afterwards. I guess I had been talking about this the whole way through + the end consolidated the experience." (AB)*

2.2.1.3.4 Expressing concern for client (1)

↑	Simon	*Session 38 HAT: "Therapist had concern at end that I may get too caught up in the issues my family and friends have" (AB)*

↑	Simon	*Session 43 HAT: "The therapist was concerned about my health" (AB)*

2.2.2 Interfering Therapist Processes (4)

2.2.2.1 Not working in client's frame of reference/acting in a controlling/directive manner (3)

↓	Joseph	*Change Interview: "sometimes, er when [T1] or [T2] goes 'oh tell me more about your mum or dad', I'm like 'oh god here we go', having to explain the whole family history, that kind of thing" (AB)*
↓	Joseph	*Change Interview: "But I get that, I mean it's not painful but it's not helpful, I mean I don't feel helpful talking about my family all the time but it's obviously helpful for them to get a context going into it" (AB)*
↓	Joseph	*Change Interview: "the delving in […] my family are my family and it's like I see them often enough, I don't have to come here and talk about them for an hour because that's not what's upsetting me, that's not why I'm here, because of my family. I feel like I have a good relationship with my family." (AR)*
↓	Sofia	*Hindering: therapist directivity (Adjudication)*
↓	Sofia	*Sofia reports struggling to come to terms with the changes she made, therefore this brief suggests that the therapist was directive and perhaps pushed Sofia too hard to discuss or explore issues she was not ready to. (AB)*
↓	Luke	*On multiple occasions, Luke's therapists wrote about creating silence in the session as an "opportunity" for Luke to engage. Luke himself did not feel like this was the therapist kindly offering him a chance to engage – he viewed it as an abandonment and a lack of caring. The second therapist wrote: "he said it was not an opportunity. He said I wasn't interested.". Luke described the silence as hindering on a couple of occasions. (AB)*
↓	Luke	*Hindering: The client acknowledged that he found the second therapist's attempts to play games hindering (Adjudication)*
↓	Luke	*Both therapists seem to have had an agenda, or a condition for their regard. That condition was "you must improve or show evidence of trying to improve". (AB)*
↓	Luke	*Hindering: The initial therapist's lack of patience with the client's process, and the manipulation implied within the threat of endings as a motivational force in getting the client to engage (Adjudication)*
↓	Luke	*An example of lack of empathy shown towards Luke was in session 15, where the therapist's notes read: "I detected that*

		several times in the session while in silence he appeared to be suppressing/ swallowing 'tears'. I suggested to him that to let these feelings out might permit us to make some progress and continue the therapy." To notice what was a rare example of emotion for Luke, and to respond to it with a push rather than empathy, could have shut Luke down. It was as though the therapist was saying, "keep going, more of this!" rather than reassuring him of the safety of the space, or acknowledging that it was hard for Luke to display this emotion. (AB)
↓	Luke	*The possibility of ending seemed more like an attempt to motivate Luke than a serious concern that he couldn't benefit from therapy. In the notes for session 14, the therapist wrote under "ideas for next time": "focusing on the possibility of ending to encourage him to engage" (see Table A8). Once there was a possibility that it had "worked", and Luke was ready to engage, therapy was offered again (though an ending was suggested again). (AB)*

2.2.2.2 Therapist doubts (3)

2.2.2.2.1 Unsure how/if helping client (2)

↑	Simon	*Session 35 therapist notes: "I wonder if I am enough- what we are doing is helpful." (AR)*
↑	Simon	*Session 39 therapist notes: "Difficulty of payment but helpful/valuable for him - worth it (hard for me to always know how)."*
↓	Luke	*When Luke did not obviously appear to be trying to get better, they both went through the cycle described in Figure B2. (AB)*

2.2.2.2.2 Therapist at a loss to know what to do (2)

↓	Joseph	*Hindering = therapist's way of relating e.g.: "She did say at one point, oh I'm not sure how I can help you coz I think she felt (pause) I don't know, it was more like I was articulating in such a way that she didn't need to jump in so therefore she felt that, (hesitation) a bit hindered" (AB)*
↓	Luke	*There was a lack of consistency in how he was approached by his therapists, who were often at a loss regarding what to do with him. (AB)*
↓	Luke	*The first therapist demonstrated this most clearly by his repeated suggestions of endings, most notably suggesting ending in the session immediately after the relational assessment. (AB)*

2.2.2.2.3 Querying their connection with client (2)

↑	Simon	*Session 6 therapist notes: "He has a particular style at the sessions where he talks about many topics and it's hard to offer him feedback or observations. This has been the case since session one. I realise now I just need to let him speak and realise that, for now, he doesn't need a great deal of input from me." (AR)*
↑	Simon	*Session 35 therapist notes: "When I ask a question, the answer doesn't seem to match. Feels very 'chat' like." (AR)*
↑	Simon	*Session 34 therapist notes: "Feel I sometimes get lost in his narrative" (AR)*
↓	Joseph	*Hindering = therapist's way of relating e.g.: "She did say at one point, oh I'm not sure how I can help you coz I think she felt (pause) I don't know, it was more like I was articulating in such a way that she didn't need to jump in so therefore she felt that, (hesitation) a bit hindered" (AB)*

2.2.3.3 Inconsistent approach (1)

↓	Luke	*The therapists both seemed to go through phases while working with Luke: initially accepting his way of being, feeling frustrated with him, unsuccessfully attempting to motivate him, withdrawing, and then re-engaging. Sometimes the therapist went through this cycle in a couple of sessions, other times it took longer. (AB)*
↓	Luke	*The second therapist demonstrated this by changing her approach week to week and eventually resorting to bringing in games. The games, explained in notes from sessions 41-43, appeared to initially bemuse then frustrate Luke, and he referred to them as hindering or unhelpful in his last change interview. (AB)*
↓	Luke	*It will have been clear to Luke that his therapists were floundering and changing their approach (AB)*
↓	Luke	*Hindering: The initial therapist's lack of patience with the client's process, and the manipulation implied within the threat of endings as a motivational force in getting the client to engage (Adjudication)*

2.3 Client in the process (8)

2.3.1 Facilitative client processes

2.3.1.1 Client commitment

2.3.1.1.1 Engagement in therapeutic process (7)

↑	Julia	*During the course of therapy Julia was fully engaged from the beginning which is indicated through quantitative and qualitative data (AB)*
↑	Linda	*Having accurate and optimistic expectations would invariably enable Linda to utilise the therapy to its fullest. This can be evidenced in her described tendency to self-process outside of the session, as well as her preparation and readiness to explore specific issues in therapy (e.g Change Interviews and HAT Forms).(AR)*
↑	James	*It is notable that James engaged fully in therapy from the start and his progress is indicated qualitatively and quantitatively. (AB)*
↑	James	*James was ready to engage in therapy and there is broad evidence that he did this (AB)*
↑	Simon	*Session 21 HAT: "Now I feel I want to move forward and now that my head and heart is clearer- make a detailed plan – and take action." (AB)*
↑	Simon	*Session 27 HAT: "I realise I spend lots of time making notes – reflecting on the experience + working to deal with issues so we can get on with the good things in life + relax + be creative. The creative process of notes etc…trying to figure stuff out can be exhausting + was thinking that I'm doing it because it worked + positive break-through." (AB)*
↑	Simon	*Session 33 HAT: "This felt good because I was able to flow a bit more than usual + used time wisely as I have with other aspects of my life." (AB)*
↑	Simon	*Session 33 HAT: "1st session with new therapist (former researcher) but was tired but happy + wanted to make the most of the last seven sessions. I found a way to talk about moving on from last therapist + have a good talk through my PQ's + managed to get new therapist up to date with where am from + at + wanting to get to." (AB)*
↓	Sofia	*Sofia was very engaged in therapy (SN)*
↓	Luke	*This was shown in his consistent attendance, his requests for appointment changes, over cancellations, and his willingness to financially invest in therapy (Adjudication)*
↓	Luke	*It is reasonable to conclude from this that since Luke never even went two weeks without requesting to return to therapy, there was at least part of him that got something from the process and did not consider it a waste of his time. Towards the end of therapy he also began to express dissatisfaction if he didn't get a full session of therapy, contesting the end of the sessio , even though his lateness was his own fault. (SB)*

↓	Luke	*The fact that he wanted to continue despite this obstacle demonstrates an investment in the process incompatible with considering it a complete waste of time. (SB)*
↓	Caitlin	*It is highly plausible to suggest that while Caitlin experienced deterioration during therapy, she was nonetheless motivated to take opportunities for self-reflection and 'pay attention' to her problems. (SB)*
↓	Caitlin	*The extensive content of these as well as other qualitative data obtained from Change Interviews support the notion that Caitlin was invested in therapy and willing to use sessions to be open and reflective. (AB)*

2.3.1.1.2 Integrating therapy and life (6)

2.3.1.1.2.1 Using therapy to connect with / work on life problem (6)

↑	Julia	*HAT session 8: "Talking about getting off my anti-depressants had me feeling really anxious about it in the session, but then we turned it around and I was reminded that I do know how to ground myself in an anxious situation." (AB)*
↑	Linda	*While she did engage in therapeutic activities out with the research clinic she brought a lot of this to therapy in order to further work on aspects of herself like her inner child (Adjudication)*
↑	James	*Session 35 HAT: "A lot of things suddenly seemed to make a lot more sense and to fit together in a more coherent way. Talking about trust in general has helped me a lot to figure out who/how to trust now as an adult and also to trust more in myself." (AB)*
↑	Simon	*Session 32 HAT: "The 'help' I had provided me with positive energy + connected with my life outside the therapy room. It gave me clarity to my experience here and I guess I do feel perfect just now." (AB)*
↑	Simon	*We felt he had a lot of extra-therapy events occurring which the sceptic case argued led to improvement, but we felt that these events were very complimentary of therapy and aided by it. A virtuous cycle (as opposed to a vicious cycle) occurred, in which Simon benefitted more from his circumstances as a result of therapy and benefitted more from his therapy as a result of his circumstances (Adjudication)*
↓	Sofia	*Sofia's discomfort at living in the UK seems to become more prevalent towards the end of therapy. The therapist notes from her final session, indicate that the client discusses her dilemma over moving home or staying in the UK. (AB)*

↓	Caitlin	*Caitlin's completed "Helpful Aspects of Therapy" (HAT) forms show that a significant amount of her time in therapy was spent exploring her experience in relationships with others and how she was experiencing a number of difficulties at the time of going to therapy within these relationships. (SB)*

2.3.1.1.2.2 Using therapy to gain distance/perspective on life problem (2)

↑	James	*Session 31 HAT: "Taking the time to be in the moment instead of constantly dealing with pressing issues." (AB)*
↑	Simon	*Session 3 HAT: "I find therapy useful to talk about my problems. I think I am going to try to put it behind me now and relax for the rest of the day/week." (AB)*
↑	Simon	*Session 16 HAT: "2nd session with therapist and was feeling pretty horrible last few days- not sleeping and problems with neighbours and missing classes and struggling to feel or think positive. Where to start? Start with the heart. I am meeting my girlfriend after this + decided to start talking about happy stuff. It gave me a break from a hard week + it has been a bit relentless + mental health issues were a bit exaggerated because of tiredness and stress. The above comment showed me that I can break from the negative into the positive and then it felt good to talk about my week." (AB)*
↑	Simon	*Session 42 HAT: "I have put my daughter's problems into perspective + a limit on how they affect me" (AB).*

2.3.1.1.2.3 Reflecting on progress in life problem (1)

↑	James	*Session 25 HAT: "My therapist noted some things where I had been assertive and also pointed out that I could feel quite proud of how I had handled myself with my mother". (AB)*
↑	James	*Session 26 HAT: "I realized that I have made some progress in integrating different aspects of my life/self into a whole. This helps me to interact with other people with less trepidation." (AB)*
↑	James	*Session 40 HAT: "It was helpful for me to see/realize the many positive possible aspects of human interaction. The world seems like a nicer and less frightening place." (AB)*

2.3.1.1.3 Working really hard in and out of therapy room (4)

↑	Julia	*In her six month follow-up change interview P103 Julia clearly reports that its work done inside and outside of therapy that helped her: "… so I think being better off now came from really*

		hard work in here but also really hard work outside of the therapy room and then just hoping….” (AR)
↑	Linda	*She was reflective and processed lots outside and inside therapy (Adjudication)*
↑	Linda	*Most of Linda’s therapist’s identified Extra-Therapy Events relate to material that Linda found helpful working through in therapy; acting as situational or experientially specific examples of broader issues that Linda was able to change. For example, in Session 9, Linda’s distressing interaction with her husband may have initiated or progressed Linda’s ability to explore and work through this ongoing issue in therapy. Additionally, in session 19, Linda’s identified visit to a spiritual healer may have supported her ability to explore the inner child within the safety of the therapeutic relationship. In her Change Interviews, Linda expresses that talking to her therapist enabled her to address necessary conversations with her sister relating to abusive childhood experiences, for example. It therefore appears that Linda’s use of therapy provided her with a basis to continue working on issues outside of therapy, and vice-versa (AR)*
↑	Linda	*Change Interview: “And then there came a point when I thought, no, maybe there is issues that I really need to bring up, release, and let go in this safe environment. And really deal with things […] I'd spoken about that with my counsellor, and that night my sister had come online, and I don't know how many sessions there had been... Maybe three or four... I don't know how many […] Maybe, if I hadn't have been speaking to my counsellor, then I wouldn't have been able to talk to my sister about it?” (AB)*
↑	Simon	*Session 27 HAT: “ I realise I spend lots of time making notes – reflecting on the experience + working to deal with issues so we can get on with the good things in life + relax + be creative. The creative process of notes etc…trying to figure stuff out can be exhausting + was thinking that I’m doing it because it worked + positive break-through but need to talk to fiancée.” (AB)*
↑	Simon	*Session 25 HAT: “At the end of the session I talked about how much the PC therapy helps me + that it continues without the therapist when I’m elsewhere.” (AB)*

2.3.1.2 Self-help / self-initiated efforts (2)

↑	Linda	*Linda’s Change Interviews also reveal a number of extra-therapy events and processes that may have contributed or facilitated Linda’s change.(SB)*

↑	James	*James was preparing to move when he began therapy and saw this move as a metaphor for a new beginning. In the initial therapy sessions, the client seemed to equate the idea of moving house and having a fresh start with his hopes of making personal change at the start of therapy. (SB)*
↑	James	*Throughout the therapeutic process, James was often focused on moving, preparing to move, or unpacking from a move. Moving, unhappiness in a home, and the relief of being in a new home seemed to symbolize actual problems in James' life. This preoccupation with moving and relief of being in a new place may have provided a substitute sense of a satisfaction and a sense of temporary relief from stress, interpersonal and otherwise. It's possible that James confused the relief of moving and the initial good feeling of being in a new environment with therapeutic progress. (SR)*

2.3.1.2 Experiencing feelings (opening up/allowing self-awareness) (6)

↑	Linda	*HAT session 14. "Finally facing my marriage break up. I have closed the lid on it for so long, but it is time I dealt with it."(AB)*
↑	James	*HAT 19: "I felt that it was safe to share/express my feelings in a more emotional way and not just in an academic way. This is quite liberating and a relief to be able to unburden some of the stuff that has been weighing me down." (AB)*
↑	James	*HAT 24: "I feel like I am getting a much better perspective and getting more in touch with my feelings." (AB)*
↑	Simon	*In the change interview when the researcher asked Simon if there was anything in the therapy that was difficult or painful to talk about but still okay to talk about-his response was to say that: "There's nothing better than getting in touch with your feelings if you've been totally numbed." (SB)*
↓	Luke	*… the ruptures can be interpreted as an improvement in Luke's ability to express emotion. Although no other emotions were displayed, this does not mean they wouldn't have if therapy had continued; bear in mind the whole process was slowed by Luke's inability to fully engage, and that Luke had a lot of bottled-up anger (SR)*
↓	Caitlin	*She explicitly highlights aspects of therapy which have been beneficial to her, such as the importance of the space being 'safe' to 'project thoughts' and 'open up'. (SB)*

2.3.1.2.1 Realising feelings (3)

↑	Julia	*HAT session 1: "Talking about it makes me realize some feelings I couldn't label before. My counsellor has a really good way of putting words to it "(AB)*
↑	Linda	*Making realisations (AB)*
↑	Linda	*HAT Session 2: "I said I feel that I hide my true self – therapist said to protect myself and I think that I am <u>scared to reveal me</u> because I don't know who me is, or if I will like her." (AB)*
↑	Linda	*HAT Session 2: "I feel I am still defensive – <u>scared to upset people</u> – I need to please me more." (AB)*
↑	Linda	*HAT session 6. "Talking about my marriage break up. I feel like it was my fault I broke up the family and yet (my ex-husband) stopped loving me first – I am confused." (AB)*
↑	Linda	*HAT Session 12: "My relationship with my daughter, I realise that I had not forgiven her for moving in with my ex-husband and his wife. I have been resenting her and in a way blaming her for not supporting me. I now realise she has been and always will." (AB)*
↑	Linda	*HAT Session 18: "I feel like I have a locked door and I am afraid to open up in case I am drowned by what is inside. It made me aware that I have not forgiven myself for past events or weaknesses." (AB)*
↑	James	*HAT 23: "Allowed my inner child to come out from hiding has been helpful for me to recognize/validate my emotional self." (AB)*

2.3.1.2.2 Working with/resolving/releasing stuck feelings (2)

↑	Linda	*HAT Session 11." I started to deal with the issues of my divorce. I have been suppressing my feeling so it is good to deal with them." (AB)*

2.3.1.2.2.1 Letting go (2)

↑	Linda	*Change Interview: "And then there came a point when I thought, no, maybe there are issues that I really need to bring up, <u>release</u>, and let go in this safe environment. And really deal with things… (AB)*
↑	James	*Session 3 HAT: "Unburdening myself with regards to talking about family trauma/incest. Helpful to speak about it so I can get things clear in my own head." (AB)*
↑	James	*Session 9 HAT: "Somehow I suddenly felt like it was ok to feel ok! It was a huge relief." (AB)*

↑	James	*Session 9a HAT: "I began to see my relationship with my mother in a new light. I feel unburdened as a result." (AB)*
↑	James	*Session 19 HAT: "I felt that it was safe to share/express my feelings in a more emotional way and not just in an academic way. This is quite liberating and a relief to be able to unburden some of the stuff that has been weighing me down." (AB)*
↑	James	*Session 30 HAT: "I feel it to be helpful to give myself permission to let go of my mother through a gradual sort of respectful detachment. It gives me more space to be myself." (AB)*
↑	James	*Session 35 HAT: "Starting to deal with the feelings of shame and guilt is the deepest most important breakthrough in the therapy so far. The realization and new understanding of these issues allows me to find ways of finally tackling/dealing with the problem." (AB)*
↑	James	*HAT 36: It has all been very helpful today and last week. I feel relieved. (AB)*

2.3.1.2.2.2 Letting in (inner child) (2)

↑	Linda	*Therapist notes Session 18: "…I said that it seemed like although on one level everything was fine there was something not fine, something that still needed to be done. The tone of the session changed completely and she told me that she visualises a tiny door with a little girl hiding behind it, who seems to be her. The little girl wants to merge with her, but she's scared that if she lets her through the door she will be washed away. She got quite upset and told me that the little girl is saying sorry, but she can't forgive her. She doesn't want to let her back in because she's weak and all the bad things are her fault." (AB)*
↑	Linda	*A specific therapy process that Linda identified as being extremely helpful was the exploration or 'letting in' of her inner child, which she described as making her feel complete. (AB)*
↑	Linda	*Session 19 HAT: "Last week we spoke of my inner child being locked out – today I let my inner child in. I have not felt complete for a while and after letting my child in – I feel complete." (AB)*
↑	James	*Session 23 HAT: "Allowed my inner child to come out from hiding has been helpful for me to recognize/validate my emotional self." (AB)*

2.3.1.2.3 Greater general awareness of feelings (2)

↑	Linda	*Change Interview: "It's like, I'm more aware of where my feelings are coming from." (AB)*
↑	Linda	*Developing awareness of, and analysing feelings. (AB)*
↑	James	*Session 12 HAT: "Came to see my current circumstances within a larger context. This helped me to gauge my reactions more appropriately" (AB)*
↑	James	*Session 29 HAT: "I found it helpful and interesting to see the relationship between feelings of guilt and the question of honesty." (AB)*

2.3.1.2.4 Valuing feelings (1)

↑	James	*Session 23 HAT: "Allowed my inner child to come out from hiding has been helpful for me to recognize/validate my emotional self." (AB)*
↑	James	*Session 28 HAT: "I suppose feeling a sense of anger informs me of what more power/strength I might have to use rather in a more positive way. It was kind of empowering." (AB)*

2.3.1.3 Working through complex situations (5)

↑	Linda	*Working through complex situations. (AB)*

2.3.1.3.1 Exploring (5)

↑	Linda	*Therapeutic relationship – therapist worked within Linda's frame of reference, <u>allowing for exploration</u> (Adjudication)*
↑	Linda	*HAT Session 16: "I talked about not attending my Granda's funeral, and feeling guilty about that. It made me look at the reasons why I still feel guilty and that maybe it was all right for me to have nice memories of him even if he was cruel to my gran." (AB)*
↓	Joseph	*Helpful: a place for the client to talk and explore to be a helpful for the client (Adjudication)*
↓	Caitlin	*An aspect which stands out in these forms regarding how Caitlin experiences therapy with both therapists is 'being able' to explore the content she brings. (SB)*
↓	Caitlin	*Session 2: Discusses triggers of anger. (SB)*
↓	Caitlin	*Session 14: Explores friendship in more depth rather than using anger which is what she usually would have done. (SB)*

2.3.1.3.1.1 Sorting thoughts more clearly (4)

↑	Linda	*Change Interview: "…the therapy is definitely made me look at things, and analyse things, and think about things. Erm... I always have thought about things, but it's like, it's now..." (AB)*
↑	James	*James was able to sort out his thoughts to think more clearly. "I'm unscrambling threads and able to see now some of the threads more clearly […] They're more in a tapestry rather than spaghetti, spaghetti is a knot or hairball I don't need that in my life […] 20 years ago all of my problems kind of meshed into one horrible feeling in my stomach and I couldn't figure out what it was because it was all my emotions were all combobulated as well but now I find I'm much more able to see things clearly and distinctly." (AB)*
↑	Simon	*Change Interview: "Therapy definitely helped because it is a process I learned to respect where coming here I can separate issues." (AB)*
↓	Joseph	*When asked to identify helpful aspects in the therapeutic process Joseph is able to offer several examples which relate predominantly to the opportunity for him to talk and share his experiences with a neutral person, that he has space and time to work through things that are in his head, that he is challenged to think differently about aspects of his process and that he feels his experience is validated. (SB)*

2.3.1.3.1.2 Articulating thoughts (3)

↑	James	*Session 1 HAT: "Some of my thoughts began to coalesce sufficiently for me to articulate. It was helpful for me to have a more accurate feeling for what is to be done and what has already been done." (AB)*
↓	Joseph	*Change Interview: "My method works for me and that was what was good with therapist 1 and therapist 2, I feel like sometimes I might be talking complete nonsense, but I get to vocalise and then I go away and think 'oh actually I do do that'" (SB)*
↓	Caitlin	*She explicitly highlights aspects of therapy which have been beneficial to her, such as the importance of the space being 'safe' to 'project thoughts' and 'open up'. (SB)*

2.3.1.3.1.3 Identifying main issues (2)

↑	James	*Session 2 HAT: "Managed to identify more or less the main different issues that I am dealing with in life". (AB)*
↓	Caitlin	*Session 6: Begins to understand the extent of her anger. (SB)*

2.3.1.3.1.4 Looking at issues in depth (2)

↑	Linda	*Looking at issues in depth. (AB)*
↓	Caitlin	*Helpful: Being able to address issues (e.g. anger) – PQ items changing (Adjudication)*

2.3.1.3.2 Making connections / understanding why things happened (3)

↑	Linda	*Change Interview: "…I'm tending to look at things and think well that's because of that. If this had happened, that had happened." (AB)*
↑	Linda	*Session 14 HAT: "I have been looking for the (why) answer to my marriage break up. I feel free." (AB)*
↑	James	*Session 8 HAT: "I seem to be able to piece things together so that I could see different stuff overlaps in my life. As stuff seemed to make more sense, I became less anxious."*
↑	James	*Session 10 HAT: "I made the connection between my mother's behaviour and some of my own dysfunction especially regarding food. I have a much clearer understanding of some of the problems I am dealing with." (AB)*
↑	James	*Session 28 HAT: "A lot of things suddenly seemed to make a lot more sense and to fit together in a more coherent way." (AB)*
↑	James	*Session 37 HAT: "There has just been progressively more understanding of how all the different issues in my life fit together." (AB)*
↑	James	*Session 39 HAT: "Hard to specify exactly as it just seems as though a lot of different bits fit together now. Piecing things together has an illuminating and empowering effect." (AB)*

2.3.1.3.2.1 Recognising what issues are in the past (1)

↑	Simon	*Session 15 HAT: "I talked freely for some time and therapist agreed that there was a lot going on. This helped me to decide to look at my life and decide that somethings are in the past." (AB)*
↑	Simon	*Session 18 HAT: "Near the start I said I was feeling more positive because I had accepted that the incredibly unhealthy feelings I associated with a representative of an organisation was in the past now and no point holding on to them." (AB)*

2.3.1.3.2.2 See current circumstances in larger context (1)

↑	James	*Session 12 HAT: "Came to see my current circumstances within a larger context. This helped me to gauge my reactions more appropriately" (AB)*

2.3.2 Interfering client processes

2.3.2.1 Discomfort with process (4)

2.3.2.1.1 Difficulty opening up (4)

2.3.2.1.1.1 Difficult and painful (3)

↓	Joseph	*Change Interview: "Although the last session I had with therapist 2 was quite challenging so when I came out feeling very depressed after it, whereas that's never really happened before, I've never come out feeling upset so, but that last session I felt quite depressed by it, not because of what she said but what we talked about" (AB)*
↓	Sofia	*There are many comments that Sofia made that indicate how much of an impact therapy was having on her. In sessions 13 and 14, Sofia wrote about how difficult she found the changes she had made during therapy: "I've been feeling a bit down lately because of all the change that's happening in my life, whether it be good or bad" and "transitions are difficult and I still haven't made peace with my changes". (AB)*
↓	Sofia	*The evidence suggests that the therapy brought up painful and difficult things for Sofia, indeed in session 8 she records that "talking about my issues made me physically tired and feeling a bit sorry for myself". This indicates that it was difficult for her to discuss her issues (AB)*
↓	Caitlin	*Mid-3 Change Interview Hindering Therapy Processes: Going through things from childhood has been difficult and painful. (AB)*
↓	Caitlin	*Mid-4 Change Interview Hindering Therapy Processes: Opening up is difficult and painful. (AB)*
↓	Caitlin	*Mid-3 and Mid-4 Change Interview Hindering Therapy Processes: Having to be emotional has been difficult. (AB)*

2.3.2.1.1.1.1 Pressure to engage? (3)

↓	Sofia	*Sofia reports struggling to come to terms with the changes she made, therefore this brief suggests that the therapist was directive and perhaps pushed Sofia too hard to discuss or explore issues she was not ready to. (AB)*
↓	Luke	*He felt great pressure to engage, as evidenced by his anxiety at the start of session 22 after an ending had been suggested again the previous session; if he did not behave in a certain way, this relationship would end. (AB)*
↓	Luke	*An email from Luke to the first therapist read "I know that [the therapy] had to end because I wasn't participating enough". (AB)*

↓	Caitlin	*If we consider Caitlin's experience of participating in research, it can be inferred from this evidence that she was passionate about the development and progression of the Person-Centred Approach. This passion can be linked to perhaps a <u>self-pressure</u> which lead to Caitlin 'having to be emotional' from wanting to engage in research and 'give something back', not due to therapist pressure.*
↓	Caitlin	*Hindering: 'Forced self-exploration' (Adjudication)*
↓	Caitlin	*Mid-3 and Mid-4 Change Interview Hindering Therapy Processes: Having to be emotional has been difficult. (AB)*
↓	Caitlin	*This suggests that Caitlin felt pressure to engage with therapy in a way which caused her to experience difficulty and pain through forced self-exploration (AR)*

2.3.2.1.1.2 Feeling exposed (2)

↓	Caitlin	*Hindering: Opening up but with no resolution – left feeling intense vulnerability (Adjudication)*
↓	Caitlin	*Mid-3 Change Interview Hindering Therapy Processes: Therapy has left her feeling vulnerable – unpacking stuff but no resolution yet. (AB)*
↓	Caitlin	*End of Therapy Change Interview Hindering Therapy Processes: Being open leaving the client vulnerable. (AB)*
↓	Luke	*Hindering: The client also seemed to respond negatively to specific therapeutic processes, such as the reflection of body language (Adjudication)*
↓	Luke	*He also states that verbally communicating his experience made him feel ingenuine, leading to a lack of content within sessions (Adjudication)*

2.3.2.1.1.3 Concern about upsetting therapist (1)

↓	Joseph	*Client's concern for making the therapist upset due to their content (Adjudication)*
↓	Joseph	*Hindering = therapist became upset by things that Joseph was saying e.g.: "and then when i upset her I was like, 'oh god, if this had carried on, would there have been more of that, would I have been having to hold back so I didn't upset my counsellor?'" (AB)*

2.3.2.1.2 Struggling to find direction (3)

↓	Joseph	*He described finding the lack of direction from the therapists difficult and that he struggled with setting his own direction in*

		the therapy, adding that he would prefer to be given goal or targets and for advice to be given (SB)
↓	Joseph	*Change Interview: "…but that was good because it made me go; what do you want me to talk about, what do you want me to say?" (AR)*
↓	Sofia	*In session 19, she also wrote about how comfortable she felt suggesting things in therapy: "the fact that I suggested a new way of dealing with my concerns - comfortable enough to feel that this is something I can place as a proposal in the therapy space [sic]". (SB)*
↓	Luke	*Luke frequently brought no content into his sessions […] This was due to the fact that he struggled knowing what to say. (SR)*

2.3.2.1.2.1 Struggling to know what to say (2)

↓	Luke	*Hindering: The silence in session seemed to be hindering in making the client feel uncomfortable (Adjudication)*
↓	Luke	*Luke frequently brought no content into his sessions […] This was due to the fact that he struggled knowing what to say. (SR)*

2.3.2.1.2.2 Can feel like a chore to find useful direction for work (1)

↓	Joseph	*Change Interview: "I've felt i enjoyed the sessions, like I really look forward to them although sometimes I come going like I've had a really good day and I turn up in a good mood sometimes it feels a bit like, it feels a bit like a chore like I'm not really sure […] when everything seems to be going ok and then when that thing happened with my personal situation I really needed it but it wasn't available so it kind of fluctuated which I mean, i sort of found it annoying and sometimes" (AB)*
↓	Joseph	*Change Interview: "…and it feels like, I'm kind of like, eyeing the clock to see how much longer I can talk about where if I was being honest, I'd go, can we just wrap this up early coz I don't, but then I know there's a whole process to go through." (AR)*

2.3.2.1.2.3 Coasting

↓	Joseph	*Change Interview: "With therapist 1, I was coasting a wee bit, it was like i was turning up and just talking and going away and not thinking about it, whereas with therapist 2 I left having to think about it a bit more" (SB)*

2.3.2.1.3 Feeling that they had to change / do therapy differently/felt pressure from therapist (2)

↓	Joseph	*Change Interview: "…when she said I don't know how we can help you, I felt like, I don't know, does that mean I'm completely unhelp-able, I'm beyond saving or I'm being very unhelpful or obtuse - that kind of thing. […] I didn't take it personally, I wasn't upset by it, but it was more like, oh maybe what I'm saying, I have to change how I say it (AR)*
↓	Luke	*[Luke] had realised that "nothing is going to happen if [he] doesn't make it happen". Concerns were raised in the affirmative case about the conditional nature of Luke's return to therapy at session 20, yet when Luke finally engaged the way the first therapist had been encouraging him to, he rated the session as "moderately helpful" on the post-session questionnaire (SR)*

2.3.2.1.4 Feeling uncared for by therapist (1)

↓	Luke	*Luke described feeling uncared for by his second therapist; for him, it was a massive challenge to participate in therapy. In his fourth change interview, Luke said that his therapist seemed indifferent to his struggle in finding something to say. Data from the RSRS shows that Luke felt increasingly misunderstood towards the end of therapy. (AB)*
↓	Luke	*There are fewer examples which suggest that Luke felt uncared for by his first therapist; simply a statement made by Luke in his 35th session suggesting that he had felt "sacked" by his first therapist when they ended. This, however, is very telling, and is discussed elsewhere in this brief. (AB)*
↓	Luke	*Towards the end of therapy, Luke began to test boundaries, attempting to lengthen sessions. Since he was feeling uncared for at that time, it was possible that he was doing this as a test to determine how much his therapist cared. He also challenged her about reading the emails he sent, accusing her of ignoring them, evidence that he was lashing out due to feeling hurt. (AB)*
↓	Luke	*The weekly data also shows a fragility in the relationship between Luke and his therapists which could have contributed to his deterioration. His tone describing hindering aspects is quite angry. (AB)*

2.3.2.2 Deference toward therapist (3)

↓	Sofia	*Hindering: possible dependency on therapist (Adjudication)*

↓	Sofia	*These two TRS items identified suggest that a power imbalance occurred, and that the therapy was not in line with person-centred theory. […] There are also comments on the HAT forms in sessions 9 to 15 regarding suggestions made by the therapist which Sofia considers important and/or helpful. This would suggest that her therapist's opinion means a lot to her and that she gives what she says a lot of thought. (AB)*
↓	Sofia	*Additionally, the TRS scores suggest that the therapist took the lead in sessions and that Sofia felt less able to disagree with or correct her. (AR)*
↓	Joseph	*There was an increase in scores on the Therapeutic Relationship Scale for client non-deference, (factor 2). At session 5 of therapy where Joseph was working with therapist 1. […]Therapist 1 was more non-directive in the way she worked compared to Therapist 2. It is reasonable to conclude that while Joseph may have felt he preferred the more structured style of therapist 2, this way of working did in fact impact on his ability to realise his own agency. (AR)*
↓	Joseph	*Hindering = Inconsistency in therapy frequency e.g.I: "and then another big gap in-between the whole lot of this so it felt like i was kind of having to deal with it on my own internally when i really needed the sessions to be, carry on, but that was coz of my own work stuff and it was the summer there was apparently no one available either, it was just the way it was"(AB)*
↓	Caitlin	*While it cannot be denied that there were significant gaps between Caitlin's appointments, particularly with her second therapist, Caitlin notes in her Mid-4 Change Interview that these inconsistencies could not be helped. (SR)*

2.3.2.3 Struggle to integrate therapy and life (3)

↓	Joseph	*A significant life event during therapy which was not supported by the therapeutic process due to the unavailability of a therapist for Joseph. (AB)*
↓	Joseph	*Change Interview: "well I feel like I had a really, like I said a really big personal trauma, I was going out with someone and it fell apart and the aftermath was quite hurtful and erm (pause) the therapy was good because it was what I needed, but like I said there was a gap so I feel like I had to process it on my own and, and since, and now I've got a bit of distance from it I feel as if, I think I'm doing ok, but, like, I'm not really sure" (AR)*
↓	Luke	*…preventing the client from integrating therapeutic processes into his life, and life processes into therapy (Adjudication)*

↓	Caitlin	*While noting that she is not taken seriously and not understood by others with whom she has relationships with, a notable aspect with explicit reference to how she engages with therapy is that she finds it: "…difficult to utilise what she has learned from therapy." This difficulty which suggests that she was not able to integrate her learning from therapy into everyday life was frustrating given her motivation to engage in therapy. (SB)*
↓	Caitlin	*Caitlin expressed at her Mid-1 Change Interview that she struggles to apply what she learns in therapy to "outside the counselling room". (SB)*

2.3.2.4 Lack of engagement (2)

↓	Joseph	*Joseph's work commitments dictated that he needed to travel for work and would often need to change or cancel appointments. He would frequently be away from the area for significant periods of time and was unable to attend sessions regularly. (SB)*
↓	Luke	*In his HAT forms, Luke expressed frustration with sitting in silence and having pointless conversations. While the affirmative case attributes the blame for this to the therapist, Luke himself actually attributes a proportion of the blame to himself. In his change interviews, Luke admits he "[hasn't] used the chance to interact and relate". (SB)*

2.4 Challenges in the process (6)

2.4.1 Change of therapist (6)

2.4.1.1 Non-interfering effect of change of therapist

2.4.1.1.1 Positive perception (4)

↑	Simon	*Session 15 HAT: "It's good to be back! I can't remember when I was last here but feel positive to meet new therapist." (AB)*

2.4.1.1.1.1 Like a new start (2)

↓	Joseph	*Change Interview: "yeah, their tactics were completely different, I mean, this is why it feels like a new start, I mean working with a new person" (AR)*
↓	Sofia	*However, the therapist notes state that "[Sofia felt] excited about starting with her new therapist and feels there has been a turning point in her life lately". (SR)*

2.4.1.1.1.2 Handling it well (2)

↑	Simon	*Session 41 HAT: "The end of this session was really important to me (as were the rest). Therapist said they leave 29th March + this gave me a deadline + <u>am handling it well</u> on a scale of 1-7, in at three with this decision. This is helpful because it helps me focus + can share this with other people that affect my life and can help me find a balance with them and myself" (AB)*
↑	James	*Session 15 HAT: "I met the new therapist for the first time. I had been a bit anxious before but am now somewhat relieved." (AB)*

2.4.1.1.2 Beneficial impact (3)

2.4.1.1.2.1 Second therapist better fit for client (3)

↑	James	*…[second therapist] appeared to have more developed counselling skills. (AB)*
↓	Joseph	*Change Interview: "…and I think therapist 2 quite good because […] she'll challenge what I say or want me to pick up on something and delve into it, whereas therapist 1 was more happy to let me talk and just respond occasionally and I'm not really sure that is what I was after. I know that the point of these sessions is that I do the talking and its, I kind of lead the conversations but therapist 2 is more happy to jump in when therapist 1 was a bit held back a bit" (SB)*
↓	Joseph	*Joseph worked with two therapists over the duration of therapy, both had very different styles of working but both were working within the person-centred model of therapy. It is clear in the narrative which emerged from the change interview that he found working with therapist 2 more helpful than therapist 1 due to their personality characteristics and what Joseph perceived he had in common with them. (SB)*
↓	Joseph	*Joseph identifies difficulties in relationship building with therapist 1. (AB)*
↓	Luke	*Helpful: Overall, the change of therapist seems to be helpful for the client (Adjudication)*

2.4.1.1.3 No impact (3)

2.4.1.1.3.1 Both/all therapists were experienced as helpful (2)

↑	James	*James' responses in the Change Record suggest that he found therapy helpful while working with both of his therapists (AB)*
↓	Joseph	*Despite these differences Joseph describes both relationships with positivity. (SB)*

2.4.1.1.3.2 Attached to process not therapist (1)

↑	Simon	*Change Interview: "…obviously there was a lot more sessions with [the third therapist], but the other two that you kinda get to a point where you're spending an hour a week in a room with someone doing absolutely amazing stuff so all you can say realistically for me is that there may be some positive attachment not the person but with the process, and em, it's a really good experience. I'd encourage anybody to go and do it because, there's no side-effects." (AB)*
↑	Simon	*In doing this the therapists ensured that the overall process of therapy that Simon was engaging on carry on as smoothly as possible. (AB)*

2.4.1.2 Interfering effects of change of therapist (4)

↓	Joseph	*Hindering: change of therapists during the therapy (Adjudication)*
↓	Joseph	*Joseph's change interview, conducted at session 10 in therapy, shows that the irregularity in sessions and the change in therapist as well as great variation in therapist style had a detrimental effect on his progress through therapy. (AB)*
↓	Sofia	*Hindering: change of therapist (Adjudication)*
↓	Sofia	*Sofia's PQ scores at session 16 increase significantly […]. This suggests that Sofia has been experiencing high levels of distress in the seven days leading up to session 16. This also coincides with a change of therapist which could be adding to Sofia's anxiety. (AB)*
↓	Sofia	*Indeed, through examination of the emails between Sofia and her therapist there is a stark contrast between the amount of contact she had with the first therapist compared to the second. Indeed, she and her first therapist seemed to frequently be in touch via email and the therapist sent Sofia relevant material on certain occasions whereas there were no emails between Sofia and the second therapist which suggests that they did not have as strong a therapeutic relationship and therefore this could have negatively affected the progress Sofia was making in therapy. (AB)*
↓	Caitlin	*Hindering: Change of therapist (Adjudication)*
↓	Caitlin	*As Caitlin's deterioration became apparent from the point of her Mid-3 change interview, an aspect of her therapy process which Caitlin explicitly states as being hindering is a change of therapist on her 26th session. This change was reported in her Mid-3 Change Interview. The change in therapists also aligns with the significant deterioration in her outcome measure scores at Mid-3. (AB)*

2.4.1.2.1.1 Unfortunate timing for client (3)
2.4.1.2.1.1 In mid-process (2)

↓	Joseph	*Hindering: Therapist Changes e.g.: "yeah and it was like, it kind of felt like 'should I even bother telling this person coz now she's going to go away feeling really upset', and she did tell me, 'I think you'll be ok' and was like a really lovely person. I was like 'thanks', but like that was very final, that was it closed and I was probably never going to see her again and now she's gone away knowing all this stuff about me, which was quite hard". (AB)*
↓	Sofia	*This evidence presented here suggests that the change of therapist may have been very difficult for Sofia. This is of particular importance if this was a time when Sofia was under a lot of pressure to make a very important decision about her future and caused an increase in her levels of distress as it will have been hard to build a strong connection with the new therapist in only five sessions. This suggests that the change of therapist was detrimental to Sofia's wellbeing and indeed coincides with a peak in her PQ scores. (AB)*

2.4.1.2.1.2 Loss of support when feeling vulnerable (1)

↓	Caitlin	*Between sessions 26-31 when she was undergoing frequent changes to her medication, Caitlin also had a change of therapist at this time (session 26). She highlights both of these changes as hindering in her Mid-3 Change Interview. These hindering factors with the additional deterioration in her Mid-3 Change Interview items; 'I feel overwhelmed by my depression' and ''Sometimes I do not feel like I get much support', suggest that the change of therapist was untimely for Caitlin and did not aid her in feeling fully supported (AR)*

2.4.1.2.2 Disruption in building relationship (2)

↓	Joseph	*It is clear from the change interview that Joseph struggled with the changes in therapist and that he found it difficult to build relationships. Given the amount of sessions that he attended (16 in total) it is not unreasonable to assume that the majority of these were taken up with either establishing the relationship or ending it. (AB)*
↓	Joseph	*Throughout the course of therapy Joseph was assigned to 3 different therapists (2 which he worked with and 1 which he only had email communication with due to incompatible appointment times) and had 2 different researchers. The data reported from his therapeutic relationship scale and the working alliance inventory, illustrate the difficulty Joseph experienced in building trust in the therapeutic relationship.*

		Evidence of these difficulties which ultimately link to Josephs disengagement from therapy can be found in his change interview data, as well as in the email trail, which clearly shows the difficulties in maintaining consistent, reliable therapeutic relationships. (AB)
↓	Sofia	*There is substantial evidence to prove a positive relationship with the first therapist, the deterioration occurs after Sofia changed therapist. It can be difficult build a connection after a transfer, especially when there have been a substantial number of sessions with the first and that a strong therapeutic relationship has been built. (AR)*
↓	Sofia	*Furthermore, Sofia only had four sessions with the second therapist, this rebuttal will argue that this was not enough time to develop a relationship and that this coincided with the deterioration of Sofia's PQ scores as well as the deterioration of how helpful the sessions were. This suggests that the change of therapist was detrimental to the therapy and have a negative effect on Sofia, even though the client may not have been consciously aware of this. (AR)*

2.4.1.2.3 Reinforcing expectation of being rejected in relationship (1)

↓	Luke	*Hindering: The forced endings with both of the therapists seemed to perpetuate the client's negative relational beliefs and expectations (Adjudication)*
↓	Luke	*Particularly considering that the first therapist told Luke, in their second session together, that he (the therapist) would not end the therapy and the choice was with Luke, it must have been incredibly damaging for Luke that his therapist ended therapy. The first therapist also wrote in his notes for the first session "this relationship cannot end quickly", acknowledging the importance of avoiding ruptures or re-traumatising Luke. (AB)*
↓	Luke	*The fact that ending due to Luke's inability to engage was so frequently discussed in the first half of therapy seems quite punitive, re-enforcing Luke's greatest weaknesses; particularly his PQ items "my relationships fall apart soon after they begin", "I'm inferior to other people", and "I have nobody to turn to". The last of these had shown a decrease by the first change interview (moving from a 7 at intake to a 4 at Mid-1 but jumped back up to a 6 after therapy had ended at session 19. (AB)*
↓	Luke	*In session 18, the notes read: "I said that I think the next session will be our last. He said that <u>he expected it to go this way</u> however when asked to elaborate he chose not to explain what he meant." […] Luke's perception was that he was "sacked" by the first therapist (see Table A8, session 35), and*

		when he talked to the second therapist about this she described him smiling. The affirmative case suggests that the smile was Luke feeling like he had been proven right; that his PQ items were true. (AB)

2.4.1.2.4 Difficulties in relationship with second therapist (1)

↓	Luke	*In fact, the TRS data suggests that overall Luke's relationship with the second therapist was worse. Losing his temper with her does not mean that he would not have done the same with the first therapist, had he done something to make Luke angry. The change in expressed emotion coincides with the change in therapist, therefore it cannot be assumed that it had anything to do with a change within Luke. It is very possible that the change in Luke was his reaction to a therapist who he perceived to be finicky and unempathetic. (AR)*

2.4.2 Delay / inconsistency in process (4)

2.4.2.1 Inconsistency of sessions (3)

↑	Simon	*there was an approximate 4 week gap between sessions 7 & 8, and between, 10 & 11 (sessions 7-11 inclusive) and between session 14 & 15, (where there was another 4 week gap). (SB)*
↓	Joseph	*Hindering: lack of consistency in the therapy sessions (Adjudication)*
↓	Caitlin	*Hindering: Inconsistencies (Adjudication)*
↓	Caitlin	*Space between first and second session makes client anxious. (AB)*
↓	Caitlin	*Periods of time between sessions. (AB)*
↓	Caitlin	*… gaps of time between Caitlin's sessions, particularly with her second therapist. Email correspondence gives an indication of valid reasoning for sessions being postponed however this does not detract from the fact that there were significant amounts of time when Caitlin was not in therapy. (AB)*
↓	Caitlin	*The importance of consistency for Caitlin is further highlighted in her Mid-2 and Mid-4 Change Interviews, as she highlights that having consistent weekly sessions are 'something to look forward to' and help her to 'feel less agitated', she also notes the lack of consistency in her session times in her Mid-4 Change Interview as hindering. (AB)*
↓	Caitlin	*With Caitlin's Change Interview data offering an indication that the lack of time-consistency between sessions was hindering to her ability to improve through therapy, her weekly PQ scores also suggest deterioration following inconsistency of this nature. The most significant period of*

		inconsistency falls between session 36-41. During this time Caitlin enquired via email about the possibility of having a therapy session as she was 'struggling'. Examination of weekly PQ data reveals covariation between this period of inconsistency and a rise in PQ scores on session dates which follow gaps during this period. (AB)
↓	Caitlin	*With the exception of session 38, which saw her PQ score drop out with clinical range (3.17) following two consecutive weeks of therapy, Caitlin's PQ scores remained consistently within clinical range during this time period. As her score went below the clinical cut off following two consecutive weeks of therapy, this suggests that Caitlin was affected by the significant gaps in her therapy. (AR)*

2.4.2.1.1 Sense of abandonment (2)

↓	Joseph	*Change Interview: "… and then when that thing happened with my personal situation I really needed it but it wasn't available so it kind of fluctuated, which I mean, I sort of found it annoying" (AB)*
↓	Caitlin	*Client asks for extra session after session 38 as she is 'struggling' but cannot be offered session. (AB)*
↓	Caitlin	*Space between first and second session makes client anxious. (AB)*
↓	Caitlin	*There were notable inconsistencies with the appointment times of Caitlin's therapy which occurred with both therapists. As can be seen from email correspondence, there is inconsistency from the beginning of therapy, with a two week gap after Caitlin's first session which results in her explicitly highlighting the difficulties she experiences when faced with uncertainty: "I would prefer if my counselling is more stable though, as my anxiety goes through the roof when I'm unsure about things." (AB)*

2.4.2.2 Took time to develop relationship (2)

2.4.2.2.1 "Rocky start" (1)

↓	Sofia	*She had a rocky start in therapy and struggled to make a connection to her therapist in the first session. (AN)*

2.4.2.2.2 Took time to develop common understanding (1)

↓	Joseph	*Change Interview: "It took me a while to get anywhere with her because she didn't seem to be, not that she was baffled by what I was saying but her personal life experience was completely different to mine, she married with children or*

		divorced with children and I don't have that kind of life style and, it felt that we were trying to get a common understanding, trying to get to know each other" (AR)

3. In the end

3.1 Ending of therapy

3.1.1 Facilitative aspects of ending (4)

3.1.1.1 Feeling ready to end (4)

↑	Julia	*HAT session 9: "Letting go of sessions is scary but being able to, feels like great progress." (AB)*
↑	Linda	*HAT Session 20: "Today was the end of our sessions – I feel it is time – I feel whole and ready to embrace life." (AB)*
↑	Linda	*The significance of this event could explain Linda's end of therapy outcome data, where her PQ and CORE scores show the minimal possible, and Strathclyde Inventory show maximum possible; indicating no distress at all on any of the measures (Table A1) (AB)*
↑	James	*Session 41 HAT: "I suppose that the process of ending was handled in a good way which was helpful for me to deal with. As previously I have felt quite ambivalent about endings. But now I feel like I am successfully moving forward onto the next stage after having accomplished a lot. Progress has been made. Very beneficial!" (AB)*
↑	Simon	*Session 47 HAT: "It felt like the natural end of therapy. I feel more independent - I talked about more of happiness + health. I feel I respect the process involved + I carry it with me + already today something shifted for me before I got here. It feels like it's part of my nature" (AB)*
↑	Simon	*Session 47 therapist notes: "Our last session. Very much sense of being ready to end." (AR)*

3.1.1.1.1 Able to continue progress made (2)

↑	James	*Session 38 HAT: "Coming up with an idea of how to reflect on my feelings by writing in a way in which could help to continue the progress that has already been made. I feel more encouraged and confident and hopeful as a result." (AB)*
↑	Simon	*Session 47 HAT: "It felt like the natural end of therapy. I feel more independent- I talked about more of happiness + health. I feel I respect the process involved + I carry it with me +*

		already today something shifted for me before I got here. It feels like it's part of my nature" (AB)

3.1.1.2 Sessions helped a lot (4)

↑	Julia	*HAT session 7: "These sessions have helped a lot, often things have greatly shifted" (AB)*
↑	Julia	*This transcription is a clear evidence of Julia attributing her change to coming into research clinic and going through her therapy. (AR)*
↑	Linda	*Therefore Linda appears to attribute these developing changes relating to this long-standing problem to therapy, where change seems to be embedded in the therapy process. (AB)*
↑	Linda	*Linda's post-therapy Change Interview provides a qualitative account of this profoundly important therapeutic process, attributing this phenomenon to the counsellor offering the opportunity to allow the inner child back in. (AB)*
↑	Linda	*…although she was in momentary crisis at the time of seeking therapy, there were still issues that she was hoping to work on and change. This is highlighted particularly in Session 9, where she experienced moderate levels of distress due to an unexpected interaction with her ex-husband. Working through her marriage break-up was repeatedly identified as being a helpful aspect of therapy, leading to change. (AR)*
↑	Simon	*Looking at Simon's descriptions of his therapeutic process in this level of detail adds weight to the credibility of the attribution of the cause of the changes being due to the therapy (AB)*
↑	Simon	*What is important about this affirmative case is that as Simon attributes all the changes he achieved to the therapy, what he talks about as being helpful in most of the forty seven sessions reflects his process of coming to understand and accept more about himself, his way of being, his relationships, and his life as it is currently. In this respect this is an argument that unless proved otherwise considers that the process of change Simon went through is due to him having the time and space made available to him to work through things in his own way. (AB)*

3.1.1.2.1 Made great progress (3)

↑	Julia	*HAT session 9: "Letting go of sessions is scary but being able to, feels like great progress." (AB)*

↑	James	*HAT 41: "… now I feel like I am successfully moving forward onto the next stage after having accomplished a lot. Progress has been made. Very beneficial!" (AB)*
↑	Simon	*Session 38 HAT: "I am at the point in my life where I have made massive progress + recovered with some issues still to be worked on + that's so good that I can now steer through the positive + that I can value the time, place, and people, and resources I have in my life. Not everyone has access to therapy" (AB)*

3.1.1.2.2 Felt really good after therapy (2)

↑	Julia	*The above transcription indicates that after therapy Julia was less anxious in social situations and where in the past she had felt anxious to read a book in class now was teaching a class without feeling anxious or having a panic moreover Julia's acknowledgement that she felt good after therapy indicates a relationship between her feeling less anxious and her therapy (AR)*
↑	Simon	*Change Interview: "I feel more centred balanced, recalibrated, all the different, positive things are there because I did say a while ago that it felt like all the positive experiences I've ever had in my life flowed through me" (AB)*

3.1.1.3 Mutual decision to end (3)

↑	Julia	*HAT session 7: "These sessions have helped a lot, often things have greatly shifted but I think I'm in a place as we discussed in the session where it is more about staying well than it is about progress/ moving forward." (AB)*
↑	Linda	*Therapist notes Session 19: "We agreed to end after our next session, and both acknowledged that something huge had happened that couldn't be explained." (AB)*
↑	Simon	*Session 39 therapist notes: "Talking about where we are, working towards 45 sessions." (AR)*

3.1.2 Interfering aspects of ending (4)

3.1.2.1 Incomplete therapeutic process (4)

↓	Joseph	*At the time when Joseph ceased therapy he did not engage in any ending or completion of therapy and did not attend for his end of therapy interviews, suggesting that he felt a further change of therapist was not helpful to him and that he felt that therapy was not helpful to him. (AB)*

↓	Sofia	*Furthermore, in the final session the data shows that Sofia was struggling to make a decision regarding where to live. (SR)*
↓	Sofia	*Therefore, at the time that therapy ended Sofia was not in a good place and seemed to get very little from her five final therapy sessions. She had not fully worked through some very important personal issues. In fact therapy had brought up a lot for her, particularly in relation to her experience of life in the UK. She had not identified what it was that she felt she was missing in the UK compared to life back home. It seems like Sofia ended her time in therapy too soon and that she had not yet completed her journey to becoming a fully functioning person. (AN)*
↓	Sofia	*In her final rating of the PQ, Sofia scored the item "I feel that I haven't fully discovered myself as an adult" a 3, although this is not a deterioration from the pre-therapy score it is the highest scoring item on the PQ and would suggest that Sofia's therapeutic and personal journey is incomplete. It seems that the client ended therapy feeling lost and indecisive about how to move forward. (AB)*
↓	Luke	*…"interrupted process", where the positive effects were yet to be clearly seen (SN)*
↓	Caitlin	*What can be taken from this is that Caitlin's lack of resolution gained from therapy <u>brought her to an unsafe place emotionally</u>. She reached a point in therapy where she felt unwilling to enter further into this feeling of vulnerability and opted to end. (AR)*

3.1.2.1.2 Client decision not to continue following therapist's decision to leave / take extended break (3)

↓	Joseph	*At the time when Joseph ceased therapy he did not engage in any ending or completion of therapy and did not attend for his end of therapy interviews, suggesting that he felt a further change of therapist was not helpful to him and that he felt that therapy was not helpful to him. (AB)*
↓	Joseph	*At the point at which Joseph left therapy he would have been allocated a fourth therapist should he have decided to remain. (AB)*
↓	Luke	*Therapist email: "I remember that I said to you that I will be back in mid-August and we said that you will get in touch to restart sessions if you want them. I said that I will keep a place for you till the start of September. Also you know that I will be looking for a job too next term, which may not come quickly but if something comes up, at that point I may have to end with my clients from the university, so I am not sure exactly*

		how long I will be available to work after the summer. I am aware that if you do not come back, that I may not see you again" (SR)
↓	Caitlin	*A further element reported by the therapist to have potentially caused distress is a further change of therapist which was disclosed to Caitlin at session 44; her therapist wrote that she "Seemed slightly disappointed about having to change counsellor in the new year." Another aspect for consideration following this disclosure at session 44 was Caitlin's weekly PQ score which rose the following week. This observation by the therapist was validated at their final session as they expressed a therapy event to be: "Client got upset about ending counselling at the research clinic. Feels safe here and believes she will be vulnerable starting counselling elsewhere" (AB)*
↓	Caitlin	*…the surprising emergence of the other change items detailed above suggest that Caitlin's reported emotional vulnerability was powerful enough to prompt her to into a 'forced ending'. By initiating an end to therapy, this suggests that Caitlin had an awareness of the impact of therapy on her at his point. (AB)*
↓	Caitlin	*Client feels process of therapy too intense to cope with. (AB)*

3.1.2.2 Did not engage in ending process (2)

↓	Joseph	*At the time when Joseph ceased therapy he did not engage in any ending or completion of therapy and did not attend for his end of therapy interviews, (AB)*
↓	Luke	*They ended therapy on a very warm note, laughing together and showing regret that they didn't end properly (SR)*

3.1.2.3 Left feeling distressed and vulnerable (2)

↓	Sofia	*Sofia's discomfort at living in the UK seems to become more prevalent towards the end of therapy. (AB)*
↓	Caitlin	*This observation by the therapist was validated at their final session as they expressed a therapy event to be: "Client got upset about ending counselling at the research clinic. Feels safe here and believes she will be vulnerable starting counselling elsewhere" (AB)*
↓	Caitlin	*Ending brings a loss of safety which client has had in therapy which is also a loss of support which she has come to feel deserving of (AB)*
↓	Caitlin	*She reached a point in therapy where she felt unwilling to enter further into this feeling of vulnerability and opted to end. (AR)*

3.2 Changes (8)

3.2.1 Facilitative changes (8)

3.2.1.1 Improvements in key difficulties that wanted to address in therapy (8)

↑	Julia	*Positive change was reported on all of the items and ranged from 1to 2 points. The drastic change is quite notable. (AB)*
↑	Linda	*She also expresses that therapy enabled her to deal with underlying and pervasive issues that she had been holding throughout her life. (AR)*
↑	Linda	*All of Linda's PQ items were rated as maximum possible (7), very considerably (6), and considerably (5) within her pre-therapy assessment, and her initial therapy session. Scores for all items dropped to 'not at all' (1) post-therapy. Scores for five items reduced by 6 points, four items by 5 points, and two by 4 points throughout Linda's course of therapy. This shows a significant change in stable problems, particularly in that 6 of Linda's 11 PQ items were stated as having been stable for between 1-10 years. (AB)*
↑	Linda	*This event correlates with a significant decrease in Linda's weekly PQ scores between session 3 and 4, with her score dropping below clinical range. It also relates to an extremely important change Linda expresses in her Mid1 (10+) Change Interview, that she claims would have been unlikely without therapy: '4. I have dealt with and put to bed various childhood issues.'. The data available thus indicates that this change, and the extra-therapy processes around it, was initiated and facilitated within therapy. (AB)*
↑	James	*James was able to: a) organize and better understand his thoughts and feelings; b) increase acceptance by self and other, c) work on issues of guilt and shame, d) decrease tension and e) become more relational. These are key areas that are highly relevant to the client's personal life experience stated in his Personal Questionnaire and to how he experiences symptoms of Aspergers. (AB)*
↑	James	*James PQ Ratings show that James experienced relief of short and long standing problems as therapy progressed. (AB)*
↑	Simon	*In Simon's Personal Questionnaire he identified ten significant problems that he wanted to address in therapy. Simon rated all ten problems as being of concern to him over the previous six to eleven months. By the end of therapy, the average PQ score of these problems demonstrated that reliable change had occurred when compared with Simon how rated them at the outset of therapy (pre: 5.6 – post: 1.8). (AB)*
↓	Joseph	*In fact Joseph's personal questionnaire data, which is representative of his own perspective and process', indicates*

		that despite a slight deterioration in scores from mid or end point of therapy overall there has been a consistent and robust improvement in how he views the aspects of his life that he would like to change through therapy, (see PQ data). (SB)
↓	Sofia	*While Sofia rated the final 5 sessions as less helpful overall relative to previous sessions, she continued to identify important of helpful events on the HAT forms and rated these as "slightly", "moderately" or even "greatly" helpful. This suggests that she was still getting something positive from these sessions. (SR)*
↓	Sofia	*Although some of the data from the mid-therapy change interview is missing, of the 10 positive changes for which we do have data, Sofia noted that three of these would have been unlikely to occur without therapy. She rated only two changes "somewhat likely" to happen without therapy. Furthermore, of the 11 positive changes identified post-therapy Sofia rated that five would have been unlikely without therapy and only one was rated "likely" to happen without therapy. (SB)*
↓	Sofia	*During her post-therapy interview Sofia rates 10 of her 13 PQ items "not at all" – indicating that these had not been bothering her in the seven days prior to this meeting. This data suggests that the majority of the issues she wished to address in therapy had improved. (SB)*
↓	Sofia	*However the evidence from the final PQ shows substantial improvement in the three issues that have been around for her for over 10 years. (SR)*
↓	Luke	*The client's ability to write about his experience articulately allowed him to communicate with his therapist to a certain extent (Adjudication)*
↓	Luke	*Early on, his therapist noted that Luke was getting something from the process, noting in session 5 "he seems to be getting something from this" and in session 7 "merely spending this one hour per week reflecting or experiencing is good for him". In session 14 the notes report that Luke "responded negatively to the suggestion that he feels obliged to attend, and unequivocally to the question that he is getting something from the process". (SB)*
↓	Caitlin	*An indication that Caitlin did experience positive change through attending therapy can be observed firstly in her improved PQ score for her item 'I have trouble controlling my anger' which was added at her Mid-2 Change Interview. This item when added was rated as having bothered Caitlin 'very considerably', and at the end of therapy this had changed to 'very little'. After disclosing the anger she feels in her first session, Caitlin appears to more fully acknowledge the extent of her anger at session 6, where she notes in her HAT form to*

		have become more aware of it. Figure C1 chronologically highlights Caitlin's process of managing her anger using data from her HAT forms and Change Interviews. Caitlin's feelings of anger are prevalent throughout her time in therapy. It is clear that through exploration of these feelings and acceptance of them from her therapists, Caitlin was able to observe changes in the anger she experiences at each Change Interview. (SB)

3.2.1.1.1 Improvement during early stage of therapy (3)

↑	Linda	*Linda's initial distress reduced significantly within the first two sessions of her therapy (AR)*
↑	James	*James' PQ scores changed most dramatically at the 10+ Change Interview, suggesting that the initial sessions, with a less skilled therapist facilitated an improvement in how he was feeling. However, it should be noted that James PQ score varied only slightly after the 10th session (3.0) as it was at the last session (3.07). (SB)*
↑	Simon	*Two out of the ten individual items were removed from the Personal Questionnaire at Simon's request within the first ten sessions. At intake both items were rated 6 by Simon which equates to a very considerable level of impact up until the point he asked for them to be removed. (AB)*
↑	Simon	*Item four was removed within the first twenty sessions, it went down by 2 points from 5 to 3 (on a scale of 1-7). (AB)*

3.2.1.2 Relationship with self (8)

3.2.1.2.1 Increased self-awareness / self-understanding (8)

↑	Julia	*Change Interview: "I have increased self-awareness and better understanding" (AB)*
↑	Julia	*HAT session 8: "Talking about getting off my anti-depressants had me feeling really anxious about it in the session, but then we turned it around and I was reminded that I do know how to ground myself in an anxious situation." (AB)*
↑	Linda	*Becoming more self-aware. (AB)*
↑	Linda	*Change Interview: "It's like, I'm more aware of where my feelings are coming from. Or where things... I'm tending to look at things and think well that's because of that. If this had happened, that had happened. And I can't blame myself." (AB)*
↑	Linda	*… shows potentially related changes to Linda's sense of self, as well as her sense of self in relation to others. These are shown to develop throughout her therapy process (as seen*

		throughout all Change Interviews), showing an increased awareness of others, as well as a growing appreciation and understanding of herself. (AB)
↑	Linda	*One of Linda's most distressing PQ Items 'I feel out of control' (as identified in her top 3 most distressing items), also relates distinctly to changes throughout therapy, with Linda developing an awareness of her own feelings and processes, which allows her to regain control of her reactions, and become stiller. This process could relate to Linda's growing awareness of the impact she has on others. (AB)*
↑	James	*James was able to sort out his thoughts to think more clearly: "I'm unscrambling threads and able to see now some of the threads more clearly […] They're more in a tapestry rather than spaghetti, spaghetti is a knot or hairball I don't need that in my life […] 20 years ago all of my problems kind of meshed into one horrible feeling in my stomach and I couldn't figure out what it was because it was all my emotions were all combobulated as well but now I find I'm much more able to see things clearly and distinctly." (AB)*
↑	James	*Session 34 HAT: "My therapist was able to make many connections today between the different threads of our discussion. Increased awareness helps me to avoid pitfalls." (AB)*
↑	James	*James gained more access to, and awareness of his emotions: "I've been able to open up my emotional world – even though emotions are awkward. The boulder that was shutting my emotions I had to get rid of the boulder then that was the first step, and then 2nd step was the whole emotional world and it's all been an ongoing process for me and I've really been pleased with how much progress I've been making. […] Yeah, I wasn't even aware actually that I was in a state of perpetual panic because I was so used to it." (AB)*
↓	Joseph	*Client appeared to be quite well functioning and self-aware (Adjudication)*
↓	Sofia	*she had self-awareness (Adjudication)*
↓	Luke	*…in Luke's second change interview, when asked to sum up what made therapy helpful, Luke stated that he had become more aware, had started to look for other areas to improve in, and had realised that "nothing is going to happen if [he] doesn't make it happen". (SR)*
↓	Caitlin	*Mid-3 Change: More aware of behaviour (SB)*
↓	Caitlin	*Being able to understand different aspects of her experience linking – self-awareness (Adjudication)*

3.2.1.2.1.1 More aware of needs and their impact (7)

↑	Julia	*HAT session 3: "Realizing that where being a little hard on myself used to motivate me, it only makes me feel worse now that the reason I'm being hard on myself are more personal" (AB)*
↑	Linda	*HAT Session 1: "I talked about a previous relationship, where the person saw me for who I am and never asked for anything else. It made me think if my behaviours were different with him and if I let him see me while I may hide from others." (AB)*
↑	Linda	*HAT Session 2: "I said I feel that I hide my true self – therapist said to protect myself and I think that I am scared to reveal me because I don't know who me is, or if I will like her. I feel I am still defensive – scared to upset people – I need to please me more." (AB)*
↑	Linda	*HAT Session 4: "I spoke of regrets and I think this made me realise how I allow myself to feel undervalued." (AB)*
↑	Linda	*HAT Session 8:" My need to feel safe or have control. It is holding me back from being happy I know I can never feel safe." (AB)*
↑	Linda	*HAT Session 12: "My relationship with my dad. I have been holding back and in a way making myself invisible." (AB)*
↑	Linda	*HAT Session 13: "I realise I deserve more and that I am worthy." (AB)*
↑	Linda	*HAT Session 13: "A male friend asked me to have an affair with him – I turned him down because I decided I have played this game before. I want a new game, one where all my needs are met." (AB)*
↑	Linda	*HAT Session 17: "I talked about a disagreement with my friend. I realised what I want from life, and what I don't want, especially from the people that call themselves friends. It made me realise that I have no control over events but I do have control over how I deal with them." (AB)*
↑	James	*Session 7 HAT: "I realized that guilt and shame are holding me back and that this guilt and shame is not rational even though they were once useful as a defence mechanism." (AB)*
↑	James	*Session 16 HAT: "I become more aware of my own subconscious techniques of avoidance. It is helpful for me to see how much I resist certain avenues of investigation or activity." (AB)*
↑	James	*Session 13 HAT: "I saw some of my habits of perception and defensiveness. Greater understanding and a little more acceptance of life." (AB)*

↑	James	*Session 20 HAT: "I realized that up until now I have been largely concerned with the logistics of how to do/feel things safely and have yet to actually do/feel them." (AB)*
↑	James	*Session 21 HAT: "It is useful to realize that things have been achieved/accomplished but I also feel embarrassed by it!" (AB)*
↑	James	*Session 29 HAT: "I found it helpful and interesting to see the relationship between feelings of guilt and the question of honesty." (AB)*
↑	Simon	*Change Interview: "I was coming here just, probably offloading a hell of a lot of negative stuff and she [the therapist] just chipped in at one point, 'I've never heard you talk about your dad before', and at this point it was really pivotal. I just went, 'wow', and it hit me emotionally, I says 'WOWW! I've been coming in here and just moaning and moaning, and then started to change the thing of, but something definitely shifted [in my way] of looking at my life." (AB)*
↑	Simon	*Session 18 HAT:" In fact since last week's stressful meeting there had been positive progress with some of my feelings. I felt I had been hanging onto negative experience as if it was evidence that would support my argument for better treatment. I wasted 6 months going over what I would say about issues based in the past - very unhealthy and tipping my balance into negativity and affecting my every thought but life is better now." (AB)*
↑	Simon	*Session 23 HAT: "I can experience life through the week but sometimes negativity overwhelms me – but something has changed, and I am going to build on that positivity." (AB)*
↑	Simon	*Change Interview: "I really like sharing with someone, being close." (AB)*
↓	Joseph	*Change Interview: "My method works for me and that was what was good with therapist 1 and therapist 2, I feel like sometimes I might be talking complete nonsense, but I get to vocalise and then I go away and think 'oh actually I do do that'" (SB)*
↓	Joseph	*Change Interview: "…out loud, I realise, 'oh wait that probably what applies to a lot of things in my life, i do things my own way and i do it because it makes me feel comfortable'" (SB)*
↓	Luke	*In Luke's second change interview, when asked to sum up what made therapy helpful, Luke stated that he had become more aware, had started to look for other areas to improve in, and had realised that "nothing is going to happen if [he] doesn't make it happen". (SR)*
↓	Caitlin	*Mid-4 Change: Client acknowledges that she is deserving of support and notes this as 'surprising'. (AB)*

3.2.1.2.2 Increased self-acceptance (6)

↑	Julia	*Therapy seemed to facilitate Julia's personal development and self-acceptance (Adjudication)*
↑	Julia	*HAT session 5: "Realizing that I have become better at allowing myself to have off moments. Better at self-compassion." (AB)*
↑	James	*Session 5 HAT: "Irealized that I don't have to understand everything in order to deal with it. It is a great relief not to put myself under so much pressure." (AB)*
↑	James	*James gained greater self-acceptance: "It's okay to be me and where and how I am […] I feel freer to be me, freer to be happy I don't, I think somehow I used to feel that I oughtn't to be happy. I had a lot of shame and guilt um which was really unwarranted." (AB)*
↑	Simon	*Session 23 HAT: "…and accept that sometimes I feel unhealthy, but I have ways of coming out of it + it passes + I've worked hard to make life better + news of studies + career change have me looking at what I can bring to that + it feels good. I'm positive about now and the future." (AB)*
↓	Sofia	*The affirmative brief also argues that Sofia experiences deterioration in her happiness regarding her life in the UK. However, she scored the item "I feel inadequate on comparison with others in the UK" not at all, suggesting that this is not a concern at the time she ends therapy (SR)*

3.2.1.2.2.1 More self-compassion (6)

↑	Julia	*HAT session 5: "Realizing that I have become better at allowing myself to have off moments. Better at self-compassion." (AB)*
↑	Linda	*Change Interview: "It's like, I'm more aware of where my feelings are coming from. Or where things... I'm tending to look at things and think well that's because of that. If this had happened, that had happened. And I can't blame myself". (AB)*
↑	James	*Session 6 HAT: Confirmation that I do not have to take responsibility for others as much. It is not necessary for me to please others all the time. (AB)*
↑	James	*Session 9 HAT: "Somehow I suddenly felt like it was ok to feel ok! It was a huge relief." (AB)*
↑	James	*Session 18 HAT: "I realized that I could feel proud of how much progress I have already made. This helps me to be more considerate, compassionate and careful with myself." (AB)*

↑	James	*Session 29 HAT: "I found it helpful and interesting to see the relationship between feelings of guilt and the question of honesty. Somehow I feel like I need to forgive myself and I am now feeling more understanding and compassionate towards myself." (AB)*
↑	Simon	*Session 34 HAT: "I realised I shouldn't beat myself up about things so much as I would end up putting more pressure on myself to be more negative + that won't help. If I try to think and act outside the box I have built up, then that might help." (AB)*
↓	Sofia	*In addition, evidence from the HAT form from session 14 indicates that she was able to "to talk about my mental and emotional processes again with love and care" which indicates that any difficulty she was having in accepting her transition had improved. (SR)*
↓	Caitlin	*Mid-1 Change: Feeling less guilty about self (SB)*

3.2.1.2.2.2 Increased appreciation of self (3)

↑	Linda	*Figure C2 shows potentially related changes to Linda's sense of self, as well as her sense of self in relation to others. These are shown to develop throughout her therapy process (as seen throughout all Change Interviews), showing an increased awareness of others, as well as a growing appreciation and understanding of herself. (AB)*
↑	James	*Session 18 HAT: "I realized that I could feel proud of how much progress I have already made. This helps me to be more considerate, compassionate and careful with myself." (AB)*
↑	Simon	*Session 32 HAT: "It was my final therapy session with current therapist + wanted to make the most of it. The 'help' I had provided me with positive energy + connected with my life outside the therapy room. It gave me clarity to my experience here and I guess I do feel perfect just now." (AB)*
↑	Simon	*Session 33 HAT: "This felt good because I was able to flow a bit more than usual + used time wisely as I have with other aspects of my life." (AB)*

3.2.1.2.3 Feeling complete / more integrated/back on track (6)

↑	Julia	*Change Interview: "I feel like I am back to me" (AB)*
↑	Julia	*Change Interview: "I feel so much better in myself" (AB)*
↑	Linda	*A specific therapy process that Linda identified as being extremely helpful was the exploration or 'letting in' of her*

		inner child, which she described as making her feel complete (AB)
↑	James	*Session 26 HAT:"I realized that I have made some progress in integrating different aspects of my life/self into a whole." (AB)*
↑	Simon	*Simon's end of therapy [changes included]: "I feel my life is back on track" (AR)*

3.2.1.2.3.1 Increased self-control / stillness / calmer (4)

↑	Linda	*One of Linda's most distressing PQ Items 'I feel out of control' (as identified in her top 3 most distressing items), also relates distinctly to changes throughout therapy, with Linda developing an awareness of her own feelings and processes, which allows her to regain control of her reactions, and become stiller. (AB)*
↑	James	*Session 8 HAT: "I seem to be able to piece things together so that I could see different stuff overlaps in my life. As stuff seemed to make more sense I became less anxious."(AB)*
↑	James	*Session 7 HAT: "I felt much calmer and relaxed and less stressed and anxious as a result of this realization." (AB)*
↑	James	*James reduced overall physical stress and tension: "I feel more relieved, or released, or light or enlightened or allieved or relieved […] I feel much calmer and less anxious. Like a big thing's been lifted […] I feel calmer and more relaxed, less anxious. I feel more contained. I feel more content. I feel more, less uncomfortable with other people […] The feeling I feel of to a large extent I've been relieved of a burden I've been carrying a weight on my shoulders, and that is now pretty much gone, so I feel relieved, released, free. […] I've spent most of my life feeling panic, I don't feel panic which is huge, huge change." (AB)*
↑	Simon	*Session 3 HAT: "I find therapy useful to talk about my problems. I think I am going to try to put it behind me now and relax for the rest of the day/week." (AB)*
↑	Simon	*Simon's end of therapy [changes included]: "I feel less anxious in life" (AR)*
↓	Caitlin	*Mid-2 Change: Controlling my anger (SB)*
↓	Caitlin	*This consideration could offer an explanation for her end of therapy change (Appendix A Table A10) that she felt 'More relaxed' (AB)*

3.2.1.2.3.2 More access to emotions (2)

↑	James	*James gained more access to, and awareness of his emotions: "I've been able to open up my emotional world – even though emotions are awkward. The boulder that was shutting my emotions I had to get rid of the boulder then that was the first step, and then 2nd step was the whole emotional world and it's all been an ongoing process for me and I've really been pleased with how much progress I've been making […] Yeah, I wasn't even aware actually that I was in a state of perpetual panic because I was so used to it." (AB)*
↓	Luke	*Extracts from therapist session notes: "he got really angry when talking about relationships"; "he got angry, telling me that I had said that before"; "he asked angrily 'why do you tell me this?'"; "he criticized me. He even criticized my understanding of his criticism of me"; "he said 'you're calling me a problem now' - looking quite angry" (SB)*

3.2.1.2.3.3 More in touch with embodied self (1)

↑	James	*James increased his ability to be embodied: " My body feels better having me around I think, its less, my body feels less anxious, less uncertain abandoned […] I'm spending more time though in my body which is a good thing cuz I used to think most of my time in some part of my head which was connected to outer space. […] My body feels less anxious, less uncertain, abandoned. I feel happier now that I'm more integrated in my body." (AB)*

3.2.1.3 Self in the world (7)

3.2.1.3.1 More empowered / motivated (6)

↑	Linda	*Session 13 HAT: "I realise I deserve more and that I am worthy." (AB)*
↑	James	*Session 30 HAT: "I feel it to be helpful to give myself permission to let go of my mother through a gradual sort of respectful detachment. It gives me more space to be myself." (AB)*
↑	James	*Session 33 HAT: "I now feel better prepared to deal with the world." (AB)*
↑	James	*Session 35 HAT: "The realization and new understanding of these issues allows me to find ways of finally tackling/dealing with the problem." (AB)*

↑	James	*Session 39 HAT: "Hard to specify exactly as it just seems as though a lot of different bits fit together now. Piecing things together has an illuminating and empowering effect." (AB)*
↑	Simon	*Session 21 HAT: "Now I feel I want to move forward and now that my head and heart is clearer- make a detailed plan – and take action." (AB)*
↑	Simon	*Session 33 HAT: "1st session with new therapist (former researcher) but was tired but happy + wanted to make the most of the last seven sessions. I found a way to talk about moving on from last therapist + have a good talk through my PQ's + managed to get new therapist up to date with where am from + at + wanting to get to." (AB)*
↑	Simon	*Simon's end of therapy [changes included]: "My motivation is better than before" (AR)*
↓	Luke	*In Luke's second change interview, when asked to sum up what made therapy helpful, Luke stated that he had become more aware, had started to look for other areas to improve in, and had realised that "nothing is going to happen if [he] doesn't make it happen". (SR)*

3.2.1.3.1.1 Increased self-confidence / self-trust / self-belief

↑	Julia	*Change Interview: "I am more confident in myself" (AB)*
↑	Julia	*Change Interview: "I am more confident in public speaking and social interaction" (AB)*
↑	James	*Session 33 HAT: "I seem to have more self-confidence now." (AB)*
↑	Simon	*Session 23 HAT: "What made this important/significant is that I believe in myself more" (AB)*
↑	Simon	*Session 25 HAT: "This session, I talked about being successful + how much difference it made to me and my loved ones. I may feel a variety of ups and downs at times – but I know I am on the right path." (AB)*
↓	Sofia	*Helpful: she gained confidence through the therapeutic alliance (Adjudication)*

3.2.1.3.2 New understanding of self in relationship (2)

↑	Linda	*Session 9 HAT: "I should have never married my husband – I wasted so many years on someone who was not capable of fulfilling my emotional needs. I feel so disconnected, frustrated and needy." (AB)*

↑	James	*Session 9a HAT: "I began to see my relationship with my mother in a new light. I feel unburdened as a result." (AB)*

3.2.1.3.2.1 Less fearful of rejection and abandonment (4)

↑	Linda	*According to Linda's PQ duration, her longest-standing problem (6-10 years) was Item 7: "I feel that I fear rejection and abandonment"). Linda's Change Interview Record showed positive changes throughout therapy that could relate to this item. (AB)*
↑	James	*Change Interview: "I've been feeling less uncomfortable with other people, and so other people are less uncomfortable with me." (AB)*

3.2.1.3.2.1.1 Able to deal with conflict in relationship (2)

↑	Simon	*Simon's end of therapy [changes included]: "I effectively deal with conflicts in close relationships" (AR)*
↓	Caitlin	*Session 43 HAT: "Feeling better equipped to deal with conflicts." (SB)*

3.2.1.3.2.2 More able to put own needs first in relationships (3)

↑	James	*Session 6 HAT: "Confirmation that I do not have to take responsibility for others as much. It is not necessary for me to please others all the time." (AB)*
↑	James	*James feels freer to put his needs first over the needs of others: " I feel a responsibility that I do feel is that seeing as I I'm here in this body in this world I've got to act in this body in this world in an appropriate way and an appropriate way for me is trying to grow or get better and primarily for myself and secondarily to help other people […] I'm much more content now to just be myself and be responsible for myself and not feel that I'm responsible for the world or society or anything like that […] I feel less responsibility for the world […] I have to put myself first, every human being every living being has to put itself first. That's the kind of basic natural thing about living, I think. It's not a question of being selfish, it's just the way that life is, and yes I would like to help other people but I have to do it in a way that is appropriate, yes." (AB)*
↑	Simon	*Session 30 HAT: "Had ended relationship because my ex-partner hurt my daughter. This was a massive turning point – my ex had crossed a line + a decision was made that freed up my mind to experience positivity with family and friends – it simplified the complex" (AB)*

3.2.1.3.2.2.1 Able to ask for needs to be met (1)

↑	James	*James was able to take responsibility for getting his own needs met: "I said is it possible to get a hug and she was happy with that and I had a hug and I guess that um, physical personal connection was really helpful." (AB)*

3.2.1.3.2.2.2 Setting limits (1)

↑	Simon	*Session 43 HAT: "I'm being realistic, she has caused me pain + that has to stop – today is the day to do that, of course I want to stay. It's almost the same with my girlfriend who had to take responsibility for her mental health + I will tell her I will leave if she doesn't take care." (AB).*

3.2.1.3.2.2.3 Able to make choices for self (1)

↓	Sofia	*With the support of the therapy, she was able to act out of her personal agency and was affirmed in validating and experiencing her ability to make choices for herself (Adjudication)*

3.2.1.3.2.3 Increased awareness of impact on/of self in relationship (2)

↑	Linda	*This process could relate to Linda's growing awareness of the impact she has on others. (AB)*
↑	Simon	*Session 38 HAT: "I reminded her that I just want to make sure they don't go into meeting with each other+ with high expectations – start shouting and + get louder + then it's who shouts the loudest+ competitive, as their stuff affects me." (AB)*

3.2.1.3.2.4 Figuring out who/how to trust (1)

↑	James	*Session 28 HAT: "Talking about trust in general has helped me a lot to figure out who/how to trust now as an adult and also to trust more in myself." (AB)*

3.2.1.3.3 Increased openness (5)

3.2.1.3.3.1 To dealing with life as it comes (3)

↑	James	*James is more able to deal with life as it comes: "It's not always changing at the right rate or direction. I'm much more content now to just be myself […] I think that it doesn't really matter what happens, I can deal with it […] It's not to say that all those things are all tied up but I think I can accept that some of those things aren't ever going to be tied up and that's ok. It leaves me*

		sort of more free to feel my actual sentiments about things." (AB)
↑	Simon	*Session 10 HAT: "I had benefit appeal a few days ago but didn't get result I wanted. The best thing that came out of it was the way I handled the bad news. The money would have been helpful but know I was relieved the appeal was over. In the last month leading up to it my health had deteriorated, and I worried I would end up in hospital. I felt relieved that I could handle the decision well. It's a sign that my mental health is better." (AB)*
↑	Simon	*Session 41 HAT: "The end of this session was really important to me (as were the rest). Therapist said they leave 29th March + this gave me a deadline + am handling it well on a scale of 1-7, in at three with this decision. This is helpful because it helps me focus + can share this with other people that affect my life and can help me find a balance with them and myself" (AB)*

3.2.1.3.3.1.1 Some things not in control (3)

↑	Linda	*Session 17 HAT: "I talked about a disagreement with my friend. I realised what I want from life, and what I don't want, especially from the people that call themselves friends. It made me realise that I have no control over events but I do have control over how I deal with them." (AB)*
↑	James	*…accept that he might not always understand the meaning of some social interactions (AN)*
↑	Simon	*Session 42 HAT: "I realised that there are some things out of my control + best to get on with studies + support + new plan + put it into action + have less distraction" (AB)*

3.2.1.3.3.2 To connecting with other people (3)

↑	James	*Session 17 HAT: "I am gradually beginning to feel like a little more connection with society/humanity is possible" (AB)*
↑	James	*Session 26 HAT: "I realized that I have made some progress in integrating different aspects of my life/self into a whole. This helps me to interact with other people with less trepidation." (AB)*
↑	James	*Session 35 HAT: "Talking about trust in general has helped me a lot to figure out who/how to trust now as an adult and also to trust more in myself." (AB)*
↑	James	*Session 40 HAT: "It was helpful for me to see/realize the many positive possible aspects of human interaction. The world seems like a nicer and less frightening place." (AB)*

↑	James	*…through this he learned how to communicate more clearly, (AN)*
↓	Caitlin	*Mid-3 Change Interview: Trying to socialise more. (AB)*

3.2.1.3.3.2.1 More accepting of others' difference (2)

↑	Simon	*Change Interview: "Maybe the relationship's quite a big thing for me because, em, maybe not everyone wants to be with someone but we all have different ways that we look at our life and I really like sharing with someone, being close." (AB)*
↓	Caitlin	*Mid-1 Change Interview: Learned to cope with others' mental illnesses (SB)*
↓	Caitlin	*End of Therapy Change Interview: More accepting of other people (SB)*

3.2.1.3.3.3 To being in the moment (2)

↑	James	*Session 31 HAT: "Taking the time to be in the moment instead of constantly dealing with pressing issues." (AB)*
↑	Simon	*Simon's end of therapy [changes] were that: I am more accepting of the past and living in the moment (AR)*

3.2.1.3.3.4 To taking risks / challenging self (1)

↑	Julia	*Change Interview: "I am open to do more scary and unpredictable things" (AB)*
↑	Julia	*Change Interview: "I am open to challenge myself more" (AB)*

3.2.1.3.4 More positive/balanced/realistic perspective on life (4)

↑	Julia	*Change Interview: "I generally have a more positive outlook on everything" (AB)*
↑	James	*Session 11 HAT: "I felt reminded of the importance of balance. It helps me to put things in context and not worry needlessly" (AB)*
↑	James	*Session 13 HAT: "I saw some of my habits of perception and defensiveness. Greater understanding and a little more acceptance of life." (AB)*
↑	James	*Session 24 HAT: "I feel like I am getting a much better perspective and getting more in touch with my feelings." (AB)*
↑	James	*Session 32 HAT: "I was able to see myself from a more holistic perspective and was consequently able to laugh at some of my problems even though they still feel difficult sometimes." (AB)*

↑	Simon	*Session 17 HAT: "I actually thought about the positive effects of therapy and it felt good to recognise again that the process of therapy can help me recognise when life gets better + I can be positive + productive with my time and energy. Having been depressed for a long time, it's great to feel that shift + make way for positivity, creativity, and relaxation." (AB)*
↑	Simon	*Session 18 HAT: "This is helping me to be more realistic + move on + look forward to life rather than fear it." (AB)*
↑	Simon	*Session 21 HAT: "It was important because I can definitely say I feel less heavy about life + more positive about life. It used to feel like there was stuff holding me back when I wanted to progress. Now I feel I want to move forward and now that my head and heart is clearer- make a detailed plan – and take action." (AB)*
↑	Simon	*Change Interview: "I was coming here just, probably offloading a hell of a lot of negative stuff and she [the therapist] just chipped in at one point, 'I've never heard you talk about your dad before', and at this point it was really pivotal. I just went, 'wow', and it hit me emotionally, I says 'WOWW! I've been coming in here and just moaning and moaning, and then started to change the thing of, but something definitely shifted [in my way] of looking at my life." (AB)*
↑	Simon	*Change Interview: "I feel more centred balanced, recalibrated, all the different, positive things are there because I did say a while ago that it felt like all the positive experiences I've ever had in my life flowed through me." (AB)*

3.2.1.3 Positive change in personal circumstances (4)

↑	Linda	*Linda's outcome data indicates that she changed substantially not throughout therapy, but within the first two sessions. Paired with the information around Linda's extra-therapy events, this change would appear to be due to a significant change in Linda's life and work circumstances. (SB)*
↑	Linda	*Table A3 shows Linda's PQ scores throughout therapy, showing an immediate decrease (from maximum possible to very little) in these 'stable' problems between session 1 and 2 when Linda's employment circumstance changed*
↑	Linda	*…clear correlation between the minor increases and decreases that occur throughout Linda's therapy after session 2, and the extra-therapy events her therapist identified.*
↑	James	*The client moved house 2 times during therapy and each time he seemed to experience relief from leaving his previous surroundings that he associated with his disorganized physical and emotional states and life struggles. (SB)*

↑	James	*With each move he expressed how his circumstances had improved, and how he had his eyes set on unpacking and moving in […] It is possible that James' improved emotional state was more a reflection of his initial happiness with his new surroundings than with therapeutic gains. (SB)*
↑	James	*Near the end of therapy he moved to a place that was nicer and gave him access to nature and the outdoors, he hadn't been there long and it is possible he was still in the honeymoon phase. (SB)*
↑	Simon	*Session 4 HAT: "The therapist told me I looked well and that she could see a change in me. This was important <u>because I had come off all my medication the previous week and I am already feeling the benefits. I am tired due to the withdrawals, but I still wanted to go ahead with the session, so I can talk about the positive changes."</u> (AB)*
↑	Simon	*Session 5 HAT: " I hadn't seen therapist for 2 weeks - and there have been a lot of changes for the better. I came off all my medication and it has been different. I feel I am in a better, healthier place. Today was a better day to be more positive about life. "(AB)*
↑	Simon	*Session 22 HAT: "The course and offer of an interview is what has been missing for the last five years. I've been unemployed with major life events and poor health and now I can see more positivity in the future + if I can continue to work on my detailed plan + take action I will achieve all my goals of better health and back to employment that is meaningful." (AB)*
↑	Simon	*Session 38 HAT: "At the start I said to therapist I'm good and remember most things, but a lot has been happening. I'm trying to be clear about where am at + and what did I talk about last time – also <u>my memory is getting better but was definitely affected by my medication.</u> We talked briefly about last time to decide where I was on my timeline of events" (AB)*
↑	Simon	*At this stage in the therapy, based on the above information it could reasonably be suggested that the change in Simon's PQ scores could be due to the impact that stopping his medication had on his outlook and energy levels both of which seemed to help him to move look forward and do things with his time out with therapy. In the change interview when Simon was asked to reflect on what aspects of his life could have impacted on the therapy, he explicitly stated that: "I do have difficulties with my health and being on a load of medication is, is a hindrance, and being on any kind of medication that sedates you because you're just in the house feeling like 'what is going on', it's quite a weird experience." (SB)*
↑	Simon	*In the change interview when the researcher asked Simon if there was anything in the therapy that was difficult or painful to*

		talk about but still okay to talk about-his response was to say that: "There's nothing better than getting in touch with your feelings if you've been totally numbed." The main argument being made in terms of Simon coming off his medication is in relation to the side effect of him feeling 'numb', and having his memory affected. If Simon had not come off his medication in the early stages of therapy, it could have been that his capacity to think clearly and be able to access some of the deeper and true feelings associated with his experience if necessary within the therapy or in his life would have been limited. The suggestion here is that while on the medication and potentially affected by its side effects it would have been less likely for Simon to be able to do so. (SB)

3.2.1.4.1 Increased resources (4)

3.2.1.4.1.1 Within relationships (3)

↑	Linda	*Session 13 HAT: "My relationships with my family – understanding and appreciating them." (AB)*
↑	Linda	*Change Interview: "I'd spoken about that with my counsellor, and that night my sister had come online, and I don't know how many sessions there had been... Maybe three or four... I don't know how many... … Ehh... Maybe, if I hadn't have been speaking to my counsellor, then I wouldn't have been able to talk to my sister about it?" (AB)*
↑	Linda	*In Linda's Change Interviews (Mid 10+ and End), she discusses the importance of having extra support from both her sister and colleague who she indicates to be counsellors. Despite this not acting as formal counselling, she stresses the value of having these additional therapeutic (SB)*
↑	Linda	*Session 19 Therapist notes: "…She then spoke to someone at work who wondered if she was keeping her 'inner child' out and this was causing her problems. She was struck by the synchronicity of this conversation after our last session" (AB)*
↑	Simon	*Session 22 HAT:" I went on to talk about being put forward for an SVQIII in childcare + promise of interview with the council + that they wanted Dads + very supportive study and placements. This has changed the way I felt about a number of things. An acquaintance of my girlfriend put me forward for this and I told a friend how much it meant it to know all these people and value these relationships." (AB)*
↓	Joseph	*Change Interview: "yeah, I don't feel like I need to cry my eyes out to someone, I don't feel like I'm at that stage, I was a bit during the summer, but I dealt with that in my own way and I had my other friends to cry to if that makes sense?" (SB)*

3.2.1.4.1.2 Within self

↑	James	*Session 28 HAT: "I suppose feeling a sense of anger informs me of what more power/strength I might have to use rather in a more positive way. It was kind of empowering." (AB)*
↑	James	*Session 32 HAT: "This gives me the strength and courage to move forward." (AB)*
↑	James	*Session 38 HAT: "Coming up with an idea of how to reflect on my feelings by writing in a way in which could help to continue the progress that has already been made. I feel more encouraged and confident and hopeful as a result." (AB)*

3.2.2 Interfering changes

3.2.2.1 Deterioration / no change in key difficulties (4)

↓	Joseph	*Considered in context of the other evidence presented here it would be a reasonable assertion that Joseph experience of therapy was indifferent. The PQ scores reflect this assertion, showing little variation from pre-therapy assessment to the end of therapy with his mean score at pre-therapy of 5 and at the end of therapy of 3.33. This reduction in score hits the reliable change interval of 1.67 (Elliott et al, 2016) exactly, suggesting only minimal reliable change if at all on the personal questionnaire. Both scores remain outside the caseness cut-off of 3.25 indicating that scores remain within the clinical range. (AB)*
↓	Luke	*Across therapy, six of Luke's eleven items increased in score, two items were the same at the intake and last change interview, and three had decreased by one point (see Table A2). The first five longstanding problems on the list, therefore the most important to Luke, all deteriorated over the course of therapy. […] Luke's mean PQ score shows an almost consistent increase (meaning a deterioration in Luke's wellbeing); there are no large fluctuations and his lowest score is right before his very first session. (AB)*
↓	Luke	*In his Mid-2 change interview (please note that this was, at the time, an end of therapy interview), Luke describes that he "felt more guilty about self", and states that it would have been "unlikely" to happen without therapy. (AB)*
↓	Luke	*Luke's mean PQ score shows an almost consistent increase (meaning a deterioration in Luke's wellbeing); there are no large fluctuations and his lowest score is right before his very first session. (AB)*
↓	Luke	*Eight out of Luke's eleven PQ items directly reference his difficulty with relationships and social situations. These*

		personal attributes may have slowed down the therapeutic process, but not caused deterioration. He may have been disengaged in his course and family life, but it was the purpose of the therapy to help with that. Had Luke been shown sufficient empathy and UPR, he most likely would have been able to engage in the therapy in a way that was natural for him. Even though Luke's behaviour could be considered a contributor to his deterioration, therapy had a very large role to play in reinforcing his PQ items and therefore Luke's behaviour, so the deterioration was a result of therapy (AR)

3.2.2.1.1 Deterioration later in therapy (2)

↓	Sofia	*Sofia's discomfort at living in the UK seems to become more prevalent towards the end of therapy. The therapist notes from her final session, indicate that the client discusses her dilemma over moving home or staying in the UK. The therapist notes that that Sofia is upset as this decision has been impacting negatively on her relationship with her partner. (AB)*
↓	Caitlin	*Caitlin identified twelve items in her Personal Questionnaire (Elliott et al, 2016) which she would address in therapy. Caitlin's significant deterioration began at her Mid 3 change interview in eight of these items, with her ratings on the instrument indicating a significant deterioration (pre-therapy score 2.3 – mid 3 score 5.25). At the end of therapy, three of these items still indicated deterioration from her post therapy rating, each of which were rated as long standing (6-10 years or more).* • *I feel put down by people* • *I feel overwhelmed by my depression* • *Sometimes I do not feel like I get much support (AB)*

3.2.2.1.2 No clearer at end of therapy (2)

↓	Joseph	*In his Working Alliance Inventory, Joseph records a decrease on items 1 and 2 which relate to his feelings about the potential for change and the role of the therapist in supporting him to look at his problems (Hatcher & Gillaspy, 2006). For both of these items Joseph's score drops from 4 (very often) to 2 (sometimes) suggesting that, rather than feeling clearer about his long-standing problems, he feels at best no different and worst as though he is unable to make a change or that he is unable to be helped. (AB)*
↓	Sofia	*The Helpful Aspects of Therapy (HAT) form from Session 20 confirms that during this session the client explored what she feels she lacks in the UK and what she likes so much about her home country. Sofia rates this exploration as moderately helpful*

		as she notes that it was "important to hear myself say those words, admitting my feelings to myself [sic]". This indicates that the client is struggling with her life in the UK and is closely related to a number of issues she has discussed throughout her time in therapy. It would appear that the client is distressed by this decision. In addition, the Client Post Session Questionnaire (Elliott, 2016) reveals that the client "did not get anywhere" in this session. This brief argues that this session did not help her to make any progress with regards to this life changing decision of whether to move back to her home country. She rated this session as "neither helpful nor hindering" (see Appendix A, Table A15) indicating that she did not get anything particularly positive or helpful from the session, other than her exploration of the differences between the two countries. This suggests that this final session did not bring her any closer to making a decision regarding the move. (AB)

3.2.2.1.3 Scores suggest did better when not in therapy

↓	Luke	*The biggest reductions (therefore improvements to Luke) are between the intake and the first session, and between the 19th and 20th sessions. As mentioned in Appendix A, an ending occurred at the 19th session, and the gap between sessions 19 and 20, at 22 days, was the longest stretch of time Luke went without a session over the course of his therapy. This suggests that Luke did better when he wasn't in therapy. (AB)*

3.2.2.2 Disappointment with process (4)

3.2.2.2.1 Feeling worse (4)

3.2.2.2.1.1 More depressed/distressed (4)

↓	Joseph	*Change Interview: "Although the last session I had with therapist 2 was quite challenging so when I came out feeling very depressed after it, whereas that's never really happened before, I've never come out feeling upset so, but that last session I felt quite depressed by it, not because of what she said but what we talked about and it kinda, I came out feeling quite heavy about it and err and I had to go and meet a friend afterwards and it kind of affected my conversation with my friend because I was having to perk up for my friend when I've just come out feeling bit down about stuff." (AB)*
↓	Caitlin	*Mid-3 Change Interview: More depressed (AB)*
↓	Caitlin	*At the end of therapy, three of these items still indicated deterioration from her post therapy rating, each of which were*

		rated as long standing (6-10 years or more), including: I feel overwhelmed by my depression (AB)
↓	Caitlin	*It is possible to reconstruct Caitlin's process of deterioration in her PQ item "I feel overwhelmed by my depression" using information from changes identified by her during change interviews and her accounts of hindering therapy processes. (AB)*

3.2.2.2.1.1.1 Less accepting of self (2)

↓	Sofia	*Sofia's scores deteriorated on nine of the 16 items on the Strathclyde Inventory between her pre and post therapy measures. In particular, there was a reduction of four points on the item "I have condemned myself for my attitudes or behaviour" and a reduction of two points on "I have lived fully in each new moment". These seem to be linked to the PQ item which scores highest on the final measure "I feel that I haven't fully discovered myself as an adult". This suggests that Sofia has deteriorated and is unhappy when she ends therapy and seems linked to the hindering aspects of therapy she identified such as: "I felt uncomfortable with myself" in session 16 as she is struggling to accept herself. (AB)*
↓	Luke	*It is likely that believing he wasn't doing well in therapy contributed to a decrease in his self-worth. It was damaging to Luke to string him along with that threat above him; that if he didn't participate enough, therapy would end. (AB)*
↓	Luke	*In his Mid-2 change interview (please note that this was, at the time, an end of therapy interview), Luke describes that he "felt more guilty about self", and states that it would have been "unlikely" to happen without therapy (AB)*

3.2.2.2.1.1.2 Struggling to make peace with changes (1)

↓	Sofia	*Additionally, the TRS scores suggest that the therapist took the lead in sessions and that Sofia felt less able to disagree with or correct her which suggests that Sofia did not have enough support from the therapist to feel comfortable with these changes and perhaps that they were not conclusions she came to of her own accord or that she was fully accepting of. This is also evidenced in the number of suggestions that the first therapist makes and that Sofia rates as helpful, she seems to accept a lot of what the therapist offers her which may influence her ability to separate changes she is conscious of and ones that the therapist points out. (AR)*
↓	Sofia	*Sofia reported making a lot of progress and changes up until session 16; however evidence suggested that she was struggling to make peace with these. She expressed feeling down and*

		emotional due to the amount of positive and negative change she was experiencing. It may have been too much for her to cope with and was perhaps not explored in the most appropriate way by her therapist, particularly if the therapist was directive. (AN)

3.2.2.2.1.2 More vulnerable (3)

↓	Sofia	*Furthermore, one of the changes that Sofia identified mid-therapy is both negative and concerning: "pulling my hair", indeed this is also something she contacts her therapist about in an email asking for help with this issue. This suggests that there was indeed deterioration in therapy which led to her harming herself. (AR)*
↓	Luke	*There is plenty of evidence that Luke struggled with his family and he was unable to express the anger that he felt towards them. This is evidenced in an email Luke wrote: "I just try not to react but it makes me really angry and every time this happens I'm making a mental sacrifice for the sake of not arguing, I've been doing this my whole life and it seems that it destroyed me." (SB)*
↓	Caitlin	*A convincing aspect of Caitlin's deterioration as a result of the therapy process which stood out in her Change Interview data was her feeling of vulnerability when she reaches significant emotional depth in therapy. Figure B1 draws upon Caitlin's Change Interview data to demonstrate the way in which the intensity of her experience in therapy contributed to her deterioration. (AB)*
↓	Caitlin	*Mid-3 Change Interview: More vulnerable (AB)*
↓	Caitlin	*Mid-3 Change Interview: More helpless (AB)*
↓	Caitlin	*Hindering: Opening up but with no resolution – left feeling intense vulnerability (Adjudication)*
↓	Caitlin	*Forced ending as things were getting quite intense and coming to a head. (AB)*

3.2.2.2.1.2.1 Less willing to engage emotionally (2)

↓	Luke	*Luke was very emotionally shut down and defeated even towards the end of therapy, writing in an email "What's the point of getting annoyed at something if later on it loses all the meaning… At first it hurts but then you forget that it's there at all and by the time you know it you let people walk all over you and use you and you just go with it because there's nowhere to run." (SR)*

↓	Caitlin	*While Caitlin's weekly PQ began to improve gradually from session 45, one may be inclined to consider that Caitlin was less motivated to engage emotionally to the extent that she had been by the end of therapy due to the intensity of it becoming too much for her to cope with. (AB)*
↓	Caitlin	*Mid-4 Change Interview: Less engaged with others (AB)*

3.2.2.2.1.2.2 Feeling less supported (1)

↓	Caitlin	*At the end of therapy, three of these items still indicated deterioration from her post therapy rating, each of which were rated as long standing (6-10 years or more): 'Sometimes I do not feel like I get much support' (AB)*

3.2.2.2.2 Perception that getting nothing from the process (2)

↓	Joseph	*Hindering = Feeling as though he was wasting time or had nothing to say e.g.: "it's just sometimes, it depends on the mood I'm in, if I come here and I've nothing to talk about I feel like I'm wasting my time and their time". (AB)*
↓	Luke	*He reported multiple times that he had got nothing from the process; for example, writing "I don't know how/if counselling can help me in any way." In an email after session and saying to his therapist in the 54th session that he had nothing to show from therapy. This indicates that Luke himself felt no benefit from therapy, and even by the end wasn't sure what it could have done. (AB)*
↓	Luke	*In his post-session questionnaires, Luke demonstrated that he experienced very little progress, only describing progress in 3 sessions. He also felt very little shift in the way he saw things, only saying he felt a shift in four sessions. (AB)*
↓	Luke	*The client also states that he is unable to remember or recount experiences due to his inability to see the point in experiencing anything. This appeared to lead the client to feel like therapy was a waste of time (Adjudication)*
↓	Luke	*His apathy filling in the HAT forms (leaving 73% all but blank) suggests that he felt all aspects of therapy, including the research, were pointless. When he described hindering aspects of therapy, he described feeling like silence was a waste of his time, as was the therapist's attempts to reflect his body language. He felt like the therapist couldn't tell what was relevant and was spending too much time on trivial matters, writing after the 38th session "The counsellor spending the whole session on something that is irrelevant is just not on. (AB)*

3.2.2.3 Loss of hope (4)

↓	Joseph	*In his Working Alliance Inventory, Joseph records a decrease on items 1 and 2 which relate to his feelings about the potential for change and the role of the therapist in supporting him to look at his problems. For both of these items Josephs score drops from 4 (very often) to 2 (sometimes) suggesting that, rather than feeling clearer about his long-standing problems, he feels at best no different and worst as though he is unable to make a change or that he is unable to be helped. (AB)*
↓	Joseph	*Joseph met with 2 therapists and was assigned to a third, which later fell through due to difficulties with availability, over the course of 16 sessions of therapy which took a total of 259 days to complete. At the time when Joseph ceased therapy he did not engage in any ending or completion of therapy and did not attend for his end of therapy interviews, suggesting that he felt a further change of therapist was not helpful to him and that he felt that therapy was not helpful to him. (AB)*
↓	Caitlin	*A further element reported by the therapist to have potentially caused distress is a further change of therapist which was disclosed to Caitlin at session 44; her therapist wrote that she "Seemed slightly disappointed about having to change counsellor in the new year." (AB)*
↓	Caitlin	*Another notable hindering aspect shared in this interview was that she had been: "…unpacking stuff but no resolution yet." (AB)*
↓	Caitlin	*The nature of Caitlin's ending suggests that she was not ready for the intensity of her emotions if she were to allow things to 'come to a head' without the comfort of feeling like there could be a resolution. (AB)*

3.2.2.3.1 No resolution (3)

↓	Sofia	*Sofia's discomfort at living in the UK seems to become more prevalent towards the end of therapy. The therapist notes from her final session indicate that the client discusses her dilemma over moving home or staying in the UK. The therapist notes that that Sofia is upset as this decision has been impacting negatively on her relationship with her partner. In addition, the therapist writes that: "She (Sofia) talked about how moving to her home country would make her feel happy and relieved and that she feels differently about herself there and has a more satisfying social life which motivates her and energises her". This indeed indicates unhappiness with her life in the UK and the lack of motivation and energy she experiences living here. (AB)*
↓	Sofia	*This suggests that Sofia is unfulfilled by life in the UK and has great difficulty in accepting British culture – the latter is*

		something she first referred herself to therapy to address. Perhaps this is an obstacle on her journey to self-actualisation. The evidence of important extra-therapy events suggests that any post-therapy deterioration is not due to therapy but due to her difficulty in making the decision of moving back to her home country and the concern regarding the consequences this will have on her personal life. (SB)
↓	Luke	*The client's family and home situation which seemed to be his only social support appeared to be hopeless (Adjudication)*
↓	Luke	*Consistently throughout therapy, Luke referred to the lack of support he received from his family; he mentioned it in all five change interviews and it was one of the items on his PQ. When he began sending emails to his therapist, his parents were the main focus: it was clear he detested them and they had no real connection. (SB)*
↓	Luke	*We also felt that the client's age and circumstances (e.g. being a young student still living in the family home) prevented him from being able to change these moderating factors (Adjudication)*
↓	Caitlin	*Hindering: <u>Opening up but with no resolution</u> – left feeling intense vulnerability (Adjudication)*
↓	Caitlin	*Another notable hindering aspect shared in this interview was that she had been: "…unpacking stuff but no resolution yet."* *This is an indication of Caitlin's expectation of the therapeutic outcome based upon how much of herself she had shared in therapy. (AB)*

3.2.2.4 Negative change in personal circumstances

↑	Simon	*As much as Simon says that he felt he handled the [welfare rights] decision well, the process had an impact on Simon as it was taking up a lot of his time, energy and was a source of stress and anxiety and that his health was suffering. It also meant that Simon was also missing some of the classes he participated in which he says that he enjoys. It can, therefore, be inferred that with all this going on there are other explanations for the negative change in PQ scores over the course of these four sessions (SB)*
↑	Simon	*The therapist considers that Simon was a bit down regarding the outcome of his welfare appeal, adding that he said that it felt like a bit of an interrogation which was a difficult process to get through. (SB)*
↓	Luke	*His stress increased towards the end of therapy as exams approached, as he had not attended lectures, and he*

		increasingly felt he had nobody to ask for lecture notes or help with studying. (SB)
↓	Caitlin	*Life events – death of father/relationships (Adjudication)*
↓	Caitlin	*In addition to noting that she had stopped her medication at Mid-3, Caitlin also experienced a significant life-event within this same time period involving the death of her father. In her communication of this to her therapist via email, she does express that his death causes her to experience physical symptoms of anxiety which prompt her to enquire about a sooner counselling appointment date. We can infer from the fact that Caitlin was unable to wait until her scheduled counselling appointment due to this life-event evoking distressing physical manifestations of anxiety, that this contributed to her changes in therapy. (SB)*
↓	Caitlin	*While it has been argued that her father's death was not expressed by Caitlin to be devastating or leave her with anguish, she did see cause to seek out support from her therapist. As already stated, Caitlin expressed in an email to her therapist that this event had caused her to exhibit physical symptoms of anxiety such as teeth grinding. This indicates that while Caitlin was not suggesting that she was grief-stricken due to his death, there was still an impact on her nonetheless. The fact that she felt it a necessary action to seek out the support of her therapist following the onset of her physical anxiety symptoms postulates that she would have become more distressed if she did not have this avenue of support (SR)*
↓	Caitlin	*It has been possible to track Caitlin's outcome measure score deteriorations in relation to frequent mention in therapist notes and Change Interviews of changes to medication which she takes for her depression. Table C2 gives a chronological account of when during therapy mention of Caitlin's medication occurred and an examination of the outcome data patterns at these points. (SB)*
↓	Caitlin	*The qualitative and quantitative evidence presented in Table C1 indicates an impact from Caitlin's medication on her wellbeing. Examination of weekly PQ scores shows, particularly between session 26-Mid-3 which saw frequent medication changes, that her scores rise considerably into clinical range. It was also at the point of Caitlin's Mid-3 Change Interview that she expressed deterioration in her changes. It is feasible for the sceptic case to confidently argue that the medication changes that Caitlin experienced can be attributed to her significant deterioration over the course of therapy. (SB)*
↓	Caitlin	*Until her 31st session, Caitlin had been undergoing frequent alterations and changes to her medication which was prescribed to help with her depression, she then stopped taking her medication around the time of her 31st session. The sceptic case*

		argues through linking Caitlin's outcome data changes with these changes as an indication that therapy was not a significant cause of these changes (SN)

3.3 Potential impact of research (8)

3.3.1 Data may not be accurate (7)

3.3.1.1 May have wanted to please therapist / researcher (6)

↑	Linda	*This could be an indication towards researcher bias in promoting higher attribution to the therapeutic process. Linda's initial researcher (Pre, Mid 10+, End Interview) was a Counselling Psychologist at the clinic, and therefore may have been unintentionally inclined to promote therapeutic attribution. Despite Linda's relational assessments with her therapist indicating high levels of client non-deference (Table A4), this ambiguity or uncertainty indicates a potential tendency towards deference to Linda's researcher. Both Linda's therapist and initial researcher were females similar in age, although Linda's researcher came from a different discipline of Counselling Psychology, and also had a different ethnic background. Speculatively, this relational difference may have had some impact on Linda's ability to provide clear and unbiased attributional opinions within her change interviews. It is also worth noting that Linda failed to mention extra-therapeutic factors such as visiting a spiritual or holistic healer to her researchers, which again may indicate a different level of transparency. (SB)*
↑	Linda	*From her change interview transcriptions, Linda appeared to have an outgoing, open relationship with her researcher, with each change interview (Mid 10+/Mid 20+) lasting well over an hour long. Linda is also happy to share many more details of her life with her researcher, rather than her therapy; indicating a more friendly or indeed therapeutic interaction. Because of the nature of this relationship, Linda may have found it difficult to negatively attribute her changes to therapy. (SB)*
↑	Linda	*These relational differences are further highlighted in Linda's 6-Month Follow Up Interview, which was carried out by a male in his forties. Changes were attributed significantly less to therapy, with two out of four rated somewhat likely without therapy, one rated somewhat unlikely, and only one rated unlikely without therapy. Both changes rated as likely without therapy (1. I do count. and 3. I am better able to ask for what I need.), relate to previously unlikely changes from Linda's Ending Interview with her previous researcher (3. I know my own worth and don't accept less (for myself).). Whilst this may be due to a longer period of retrospection, it could also support the argument that*

		Linda was more likely to attribute post and mid therapy changes to the therapy with her initial researcher. (SB)
↑	Linda	*Although Linda's Change Interview Record indicates an attributional tendency towards the unlikelihood of change without therapy, Linda's Change Interview recordings show a level of uncertainty that that isn't explored thoroughly by the researcher conducting the interview. There is therefore a disparity between the certainty reflected in the Change Interview Record, and the actual uncertainty expressed by Linda within the interview. (SB)*
↑	James	*James may have over represented his progress in therapy as well as the positive impact of therapy in order to please his researcher and fulfil his perfectionistic needs. (SB)*
↑	Simon	*It could also be suggested that Simon was possibly being generous with his scoring, in terms of the H.A.T forms he rated every aspect as a 9, for extremely helpful, and reported no hindering aspects whatsoever about the therapy. Likewise, with the Therapeutic Relationship Scale (Carrick, 2013), with the exception of the first instance in which the questionnaire was completed (Session 4), he rated the factors of quality of the relationship and client non-deference as high in all the subsequent instances of measurement. (SR)*
↓	Joseph	*Throughout his change interview Joseph consistently alluded to his discomfort in giving any critical feedback in relation to his therapists or therapy. This is reflective of a key aspect of his personal questionnaire where he feels that he cares too much about what people think. (AB)*
↓	Sofia	*During her mid-therapy change interview, Sofia's CORE-OM score is within the clinical range but she identifies a number of positive changes. How can we account for these inconsistencies? The emails between Sofia and her therapist show a positive relationship between the two and suggest that Sofia thinks very highly of her opinion. Therefore, it may be that Sofia felt it necessary to identify as many changes as possible to justify the therapy. (AR)*
↓	Caitlin	*In specifically wishing to aid in research of the person-centred approach, it could be argued that her affinity to this particular approach is a result of her perceived positive experience of person-centred therapy. (SB)*

3.3.1.2 Hard to capture experience (5)

↓	Joseph	*Change Interview: "I can see, from an academic reason why you want to know this stuff and, but every time, but sometimes I'm like I don't really know what this question is wanting from me right now, it's like, at what point did this revelation occur, was it*

		helpful was it hindering and it's like, it's in the middle, so sometimes I'm just filling it in and ticking it out so I'm not handing back a blank form" (SR)
↓	Joseph	*Joseph has previously indicated that he struggles to provide an accurate response to the measures administered. (SB)*
↓	Joseph	*When asked about the usefulness of the research process to his therapy experience Joseph replied: "yea, yea and erm, another one, there's another form you get at the end where you have to go, at what point did this thing happen, at what point was helpful and it's really hard to go, well did it happen at 23 minutes past or did it happen at quarter past, and then you have to go was it the middle of the session and then like was helpful or harmful and I'm like it was neither, we just spoke about it, I mean sometimes we just spoke about it, I mean sometimes there is a revelation". Adding: "yea, but you're asked to put into words on a form and like you're just want to write i don't know (congruent laughing) but i mean, I'll write it and fill it in but like sometimes, not like it feels like a chore but like, it's hard to summarize what you've just spoken about for an hour into a few sentences that helps someone fill in on a data base". These excerpts highlight the difficulty in being able to gain a reliable picture from the data without considering all aspects and as such the conclusion that Joseph made a substantial deterioration over the course of therapy and that this deterioration was substantially due to the therapy is unreliable. (SB)*

3.3.1.2.1 Hard to shift between forms and session (2)

↑	Julia	*Change Interview: "The end of the interview questionnaire or the end of session questionnaires. I remember sometimes thinking that they were hard to fill out because we just had an hour of intense talking about feelings and I couldn't even remember what we started talking about sometimes." Here Julia reports how after the end of session she wouldn't remember most of her experiences in session and it can be argued that her positive score might be artificially inflated resulting in unreliable data. (SR)*
↓	Joseph	*Change Interview: "but yeah its fine, I don't mind filling out the forms but then sometimes it feels like, tick tick tick tick tick, jump into it, that kind of thing" (SR)*

3.3.1.2.2 Sometimes tired / unwell (2)

↑	Simon	*Change Interview: "I knew it was a complete package, so I was quite accepting of it, there was sometimes where I was too tired, too emotional or too numb, to be doing it, but I only left it*

		out a few times. I knew it was all part and parcel of what was going on here." It also a partial part of his experience that there were some occasions in which he felt that doing some of the forms were too much for him for various different reasons. The forms he didn't complete, however, was not the PQ outcome measure, it was sometimes that he didn't answer certain questions in the SI and Core OM measures, and after some sessions didn't do at all or didn't answer all the questions on the H.A.T. forms. (SB)
↓	Sofia	*However, the rebuttal argues that this is not the case particularly in session 16 when the scores do show a quantifiable deterioration as she specifically states feeling unwell and writes that the therapist being supportive was helpful. This evidence indicates that she was feeling uncomfortable due to her menstrual cycle and to feeling unwell rather than to any in-therapy. This is something she also mentioned in session 6: "I haven't been feeling well due to illness so I was a bit more emotional than usual" which indicates that any negative feelings in or after a therapy session tend to be linked to an extra-therapy event for Sofia, such as her health. (SR)*

3.3.1.2.3 No adjustments made (1)

↑	James	*It is likely that aspects of AS may have impacted the results of some the quantitative data as well, all which has been taken into account in this Affirmative Brief. (AB)*
↑	James	*As has been discussed previously in this case, individuals with AS have difficulty discerning and understanding social interactions. This is true for James in addition to his reporting that he is a very slow processor of information. As a participant in the Research Clinic's study, James was asked to complete questionnaires with complex relational questions in a short period of time which may have been difficult for him to do. While it is known that individuals with Autism Spectrum Disorders are at a higher risk of developing social anxiety due to cognitive and relational assessment deficits, it does not appear that any specific accommodations were made to take these processing issues into account for him as a research participant, aside from the researcher telling him to feel free to take his time. (AR)*
↑	James	*It is clear in this case that James did his best to communicate his experience of therapy using the skills he has in speaking and relating, given the stress and overwhelm he can feel when he has to fill out forms and not being able to address one thing at a time. The contradictions across data in James' case can only be analyzed in the context of the limitations and impact of his Asperger's. As is noted above, James is very clear and precise in his spoken language and he strives to convey his experience*

		clearly and correctly. He noted that this skill, and the opportunity to practice communication with his therapist, improved his ability to communicate. (AR)

3.3.1.3 Self-conscious (3)

↑	Julia	*Another aspect to be considered is that during her six month follow-up change interview Julia relates being in research to being in therapy: " yeah, maybe. Definitely. It didn't feel like I was doing a research program, it just felt like I was talking to a therapist but at the same time I knew that it was going to like because there's one (inaudible) in front of you at all times so I think you just maybe I tried to think a little more about my answers before I blurted out something just to make it more coherent." (SR)*
↓	Joseph	*When asked about aspects of the research which have been hindering to the therapeutic process, Joseph states: "yeah the one thing, what I did kind of feel, I'm not sure, there seems to be one email address for everyone and everyone relays from the same email address so I'd email therapist 1 but I'm like, am I emailing therapist 1 or is someone else going to read this, that kind of thing, or like if was emailing you to day I'm like saying hi (detail removed for confidentiality) but anyone could be picking this up, it just felt like it was one email for everyone so (faltered speech)" (AB)*
↓	Sofia	*Social desirability refers to the individual's desire to be socially acceptable when completing questionnaires (Richman, Weisband, Kiesler & Drasgow, 1999) this research suggests that the relational data may not be without limitations due to Sofia's issues with self-confidence and social identity which suggest a desire to be accepted. This is an issue she mentioned upon her referral to the Research Clinic and in items she included in the PQ: "I don't have a sense of social identity" and "I occasionally feel smaller than other people" (AR)*

3.3.1.4 Amount of paperwork hindering (1)

↓	Luke	*Luke did not mention this unprompted, but when asked in the change interviews he responded, in three of the five interviews, that he found the amount of paperwork hindering. Although this in and of itself may not have been a large factor in Luke's deterioration, it might have contributed to his overall feeling that he was wasting his time. (AB)*

3.3.2 Increased self-worth (3)

↑	James	*Being part of the research study gave the client a sense of value and relevance he might not have otherwise felt. The client frequently shared that he felt a strong sense of importance, value and giving back as a participant in the research study. His comments seemed somewhat extensive in this regard, and it is possible that the attention he received from the research during the change interviews gave him a sense of value he might not have otherwise received. Likewise, it is possible that being involved in the research study provided a temporary source of esteem he wouldn't have otherwise received. (SB)*

3.3.2.1 By helping others (3)

↑	James	*Change Interview: "It was great to see students learning." (SB)*
↑	James	*Change Interview: "It's a huge potential benefit to many, many people and I think its brilliant to be doing that." (SB)*
↑	James	*Change Interview: "It's been really good to be involved in the research. I've gained a lot for myself. Primarily I'm happy for myself and I do like to help other people if I can." (SB)*
↑	Simon	*Change Interview: "I kinda knew the deal when I was going into it, I knew that it may benefit others" (SB)*
↓	Caitlin	*Qualitative evidence extracted from Change Interviews explicitly highlights Caitlin's positive reaction to being involved in research as a participant. Furthermore evidence to support this reaction can be drawn from every Change Interview that Caitlin participated in, suggesting that this was a key motivation for her in attending therapy. The qualitative data indicates that Caitlin experienced a strong sense of altruism through her participation in the research process, particularly through helping to progress research in the person-centred approach. (SB)*

3.3.2.2 By helping self (3)

↑	James	*Change Interview: "I found the research quite gratifying. I feel less burdened by it. I don't feel like I'm taking resources. It's good for me to know you're getting benefit." (SB)*
↑	James	*Change Interview: "I do like to help and I'm glad you're doing it, cuz it's certainly what is needed most in the world. I don't want to take more than my fair share of the world and I don't want to be greedy and deprive others." (SB)*

↑	James	*Change Interview: "It's been really good to be involved in the research. I've gained a lot for myself. Primarily I'm happy for myself." (SB)*
↑	Simon	*Change Interview: "I'm quite happy that I had these sessions free of charge for all this time. It's made a massive difference to me." (SB)*
↓	Caitlin	*Furthermore, by being a participant and 'giving something back', we can infer from this that there may have been an expectation to gain something for herself through therapy – such as a resolution for the difficulties she was experiencing at the time, which has been previously discussed earlier in this rebuttal. (AR)*

3.3.3 PQ provided motivation & focus (2)

↑	Simon	*Session 4 HAT: "I was able to reduce my PQ scores and I feel good about being able to remove some soon." (AB)*
↑	Simon	*Session 5 HAT:" I can't talk about my PQ's anywhere else as it's specific to counselling. I have a good formal and informal support network but it's not always possible to talk about various aspects of health. <u>It was good to know how far I've come on in two weeks</u>." (AB)*
↑	Simon	*Session 33 HAT: "This felt good because I was able to flow a bit more than usual + used time wisely as I have with other aspects of my life. Main aim is to use time wisely + be independent in seven sessions + <u>try and clear the PQ's</u>." (AB)*
↑	Simon	*When Simon was asked in the final Change Interview what, if anything, had been helpful about taking part in the research, for example completing the PQs, Simon's response was: "Eh, I think being able to focus on something that's positive, and to be able to put it in its place. To actually use positive words and if you felt bad you'd go for that experience but ultimately it felt really good to...phew...it's like part of the motivational thing. It's a confirmation of the positivity and... I mean it's part of the 'try harder' thing that even though I'm shattered and all that just, it puts things in its place and also that at some point you have to take responsibility for your mental health, for your life for all these different things that have brought you here and it's a bit of responsibility that you've been asked to be involved in this so I'm very positive about that." (AB)*
↑	Simon	*It could have been taking part in the research that motivated Simon to keep going in terms of working on improving his PQ items. Also, Simon's sense of altruism could be a motivational factor as he knows that by demonstrating change through completing various different outcome measures all of which in*

		Simon's view support his change process and for him serve as evidence that the change must have been due to therapy (SB)
↓	Caitlin	*Regarding her use of monitoring tools, she expresses the personal benefit she experienced in being able to view her feelings from previous weeks in retrospect. She also noted in her end of therapy interview that completing a weekly PQ gave her a way to 'pay attention' to her problems. (SB)*