Explaining and Predicting Psychological Problems: The Joint Importance of Positive and Negative Constructs

A thesis submitted to the University of Stirling for the degree of
Doctor of Philosophy: PhD in Clinical Psychology

October, 2017

Andrew (Andy) Philip Siddaway

DECLARATION

In accordance with the Regulations for Higher Degrees by Research, I hereby declare that the whole thesis now submitted for the candidature of Doctor of Philosophy is a result of my own research and independent work except where reference is made to published literature. I also hereby certify that the work embodied in this thesis has not already been submitted in any substance for any degree and is not being concurrently submitted in candidature for any degree from any other institute of higher learning. I am responsible for any errors and omissions present in the thesis.

Candidate: 

Andrew Philip Siddaway

LIST OF CONTENTS

LIST OF TABLES

Chapter 5

Appendix

LIST OF FIGURES

ABSTRACT

Positive Clinical Psychology (PCP) argues that positive and negative psychological constructs are jointly important for explaining psychological problems. “Positive” constructs have been explicitly focused on by positive psychology researchers and “negative” constructs have been explicitly focused on by mental health researchers. This thesis examines the relationship between positive and negative constructs in relation to four psychological problems: depressive symptoms (Chapter 2), anxiety-problems (Chapter 3), suicide attempts (SAs) (Chapter 4 and 5), and nonsuicidal self-injury (NSSI) (Chapter 4 and 5). Clarifying how psychological problems are most appropriately conceptualised has implications for definitions, diagnostic criteria, measurement, and clinical interventions. This thesis provides evidence that some constructs form bipolar continua, having a positive pole and a negative pole, whilst other constructs do not. Chapters 2 and 3 demonstrate that well-being and calmness respectively form continua with depressive and anxiety symptoms. In contrast, Chapters 4 and 5 demonstrate that SA and NSSI cognitions do not form a continuum with another construct. Results indicate that positive and negative constructs appear to have different relationships to one-another depending on the construct under investigation. Constructs that are common in the general population – such as depressive symptoms, anxiety symptoms, well-being symptoms, and calmness symptoms – appear to be bipolar, having a positive and a negative pole. Psychological constructs that are rare in the general population and which specifically characterise psychological problems (rather than being an extreme manifestation of a common psychological experience) – such as SA and NSSI cognitions – appear to be unipolar. The replication of scientific findings also features strongly throughout this thesis. Each chapter may therefore have a timely bearing on the emerging “replication crisis” literature.

ACKNOWLEDGEMENTS

I am immensely grateful to several people, without whom this PhD would not have been possible. Enormous thanks and gratitude go to my principal supervisor, Professor Alex M. Wood, who taught me most of what I have learned about research. I have learned a massive amount from you. I have benefitted enormously from your prodigious research knowledge and perceptive approach to publishing. The knowledge you equipped me with before this PhD meant that although I had to learn a huge amount over the last three years, I was able to hit the ground running and be productive throughout the PhD (with the usual ups and downs). I have also really appreciated the career mentoring you have provided and you spelling out the culture of academia to me when needed. As you know, academics and academic work culture have often been quite different to what I am used to in the NHS and you very helpfully and skilfully explicitly named what was going on so that I could learn and benefit.

Many thanks to my second supervisors, Professors Ronan E. O'Carroll and Rory C. O'Connor. Both of you have been friendly, encouraging, knowledgeable, diplomatic, and helpful, and it has been a pleasure getting to know you and working with you. Rory, you have a masterful knowledge of the self-injurious behaviour literature and I have really enjoyed and learned from working with you. I have really benefitted from the fact that we have disagreed at times about how to conceptualise self-injurious behaviour; ever the scientist, you were always open to and encouraging of me thinking differently. I hope I get the funding and can fully join your SBRL and that we can continue to work together.

This PhD was funded by a predoctoral MRC Clinical Research Training Fellowship entitled "Identifying and measuring key cognitions in suicidal and non-suicidal self-harm." I feel incredibly fortunate to have been awarded this funding, which covered research costs and attendance on a wide range of really useful training courses. It would not have been possible for me to collect several large datasets of SIB data, or to have learned as much as I have, without my MRC Fellowship.

I am indebted to the thousands of participants who generously gave up their time to take part in my research on self-injurious behaviour. This research would not have been possible without you.

I am also grateful to Prof Liam Delaney, who co-directed the Stirling Behavioural Science Centre until recently. I learned a great deal from Centre meetings and seminars because of the thriving learning culture you created, which encouraged open and thoughtful discussions and questioning.

Finally, massive thanks must go to my long-suffering wife, who has put up with the various ways my stress manifests. I am incredibly fortunate to have such a loving, supportive, kind, and beautiful person in my life, every day. You are the most perceptive person I know; thank you for helping me fine-tune my thinking on SIB. I will forever be grateful that you gave me George.

PUBLICATIONS ARISING FROM THIS THESIS

Chapter 2: Siddaway, A. P., Wood, A. M., & Taylor P. J. (2017). The Centre for Epidemiologic Studies-Depression (CES-D) scale measures a continuum from well-being to depression: Testing two key predictions of Positive Clinical Psychology. *Journal of Affective Disorders, 213,* 180-186.

Chapter 3: Siddaway, A. P., Taylor, P. J., & Wood, A. M. (2017). Re-conceptualizing anxiety as a continuum that ranges from high calmness to high anxiety: The joint importance of reducing distress and increasing well-being. *Journal of Personality and Social Psychology*.

CHAPTER 1

1. Introduction

1.1. The Conceptual and Methodological Context to this Thesis

This thesis is conceptually underpinned by Positive Clinical Psychology (PCP), which has advocated the joint importance of "negative" and "positive" psychological constructs for understanding and treating psychological problems (Johnson & Wood, 2017; Joseph & Wood, 2010; Wood & Johnson, 2016; Wood & Tarrier, 2010). "Negative" psychological constructs typify psychological problems; these are variables that particularly characterise psychological problems, which have typically been examined by mental health researchers. In contrast, "positive" psychological constructs are variables that have been examined by positive psychology researchers (and others, as discussed below). PCP argues that positive and negative psychological constructs are *equally* important for explaining psychological problems and therefore that it is important and beneficial for the emerging positive psychology research field to be integrated with the voluminous evidence base concerned with understanding and treating psychological problems (Johnson & Wood, 2017; Joseph & Wood, 2010; Wood & Johnson, 2016; Wood & Tarrier, 2010).

PCP puts forward two major predictions to support its argument for the joint importance of "negative" and "positive" psychological constructs when understanding and treating psychological problems. The first is that *most psychological constructs are bipolar, consisting of a positive pole and a negative pole.* For example, depressive symptoms are generally understood to be a unipolar, continuous construct that ranges from infrequent, weak depressive symptoms to frequent, strong depressive symptoms (Joseph & Wood, 2010). Conceptualising depression as a bipolar construct involves hypothesising that depressive symptoms constitute one pole of a continuum that has an opposite pole. In Chapter 2, a conceptual and methodological argument is put forward that depressive symptoms form one end of a depression-well-being continuum, where depression can be understood as the "negative" pole of this continuum, and well-being as the opposite, "positive" pole of the same continuum (Joseph & Wood, 2010). The depression-well-being continuum is hypothesised to range from high depressive symptoms, through a hypothetical zero point, to high well-being symptoms.

The second major prediction of PCP is that *"positive constructs" provide important unshared explanatory power in relation to psychological problems* (Johnson & Wood,

2017; Joseph & Wood, 2010; Wood & Johnson, 2016; Wood & Tarrier, 2010). As stated, "positive" constructs are variables that particularly characterise well-being, which have typically been examined by positive psychology researchers. Both of PCP's core predictions are explained further and explored below.

This thesis tests PCP's core predictions in relation to four psychological problems: depressive symptoms (Chapter 2), anxiety-problems (Chapter 3), suicide attempts (SAs) (Chapter 4 and 5), and nonsuicidal self-injury (NSSI) (Chapter 4 and 5). Although this thesis is predominantly contextualised by PCP, it is also, to a lesser extent, contextualised by the methodological need for scientific findings to replicate, which has a timely bearing on the current "replication crisis." The replication crisis is an emerging literature debating whether important scientific findings replicate and the implications of this.

An overview of PCP and the replication crisis are provided next, along with details of how this thesis positions itself and has relevance for each of these literatures. Before this discussion, the broader context for focusing equally on positive and negative constructs is briefly outlined and "positive" and "negative" are further defined. Importantly, and as discussed further below, PCP has not been explicit or consistent in stating and explaining its core tenants, which may have led to some misconceptions about its conceptual position and aims (Held, 2016). This thesis attempts to make a contribution in its own right by clarifying and testing PCP's core messages and making them more explicit.

1.2. The Joint Importance of Positive and Negative Constructs in Explaining and Predicting Psychological Problems

PCP draws attention to the joint importance of positive and negative psychological constructs for explaining human suffering (Johnson & Wood, 2017; Joseph & Wood, 2010; Wood & Johnson, 2016; Wood & Tarrier, 2010). This, of course, is not a new claim. There are many indications in popular culture and psychological science that human existence can be understood in terms of positive and negative facets, and that positive and negative are complementary (rather than opposing) forces. In Chinese philosophy, yin and yang – translated as "dark" and "bright" or "negative" and "positive" – describe how seemingly contradictory forces are actually interconnected and interdependent in the natural world, and how they may give rise to each other as they relate to one another. Everything has yin and yang (for instance, shadow cannot exist without light) and the taijitu symbol that represents yin and yang shows a balance between two opposites with a portion of the opposite element in each section.

Different religions often refer to positive and negative as a means to make sense of and explain the phenomenology of human existence (e.g., God and the Devil). Comic

books almost always involve the interplay of heroes and villains to denote the good and bad facets of humanity, often exploring how people, even heroes, are neither totally "good" nor totally "bad." In emotion research, positive and negative affect are both recognised as important facets of human existence that have fundamental and equally important relationships with behaviour, motivation, and cognition (e.g., Barrett & Russell, 1998; Carver & Scheier, 1998; Russell & Carroll, 1999a; Yik, Russell & Steiger, 2011).

Positive and negative constructs have been examined in relation to psychological problems, and this thesis particularly tests PCP's conceptualisation of the relationship between positive and negative constructs (see Johnson & Wood, 2017; Joseph & Wood, 2010; Wood & Johnson, 2016; Wood & Tarrier, 2010). Many scholars outside of PCP have also drawn attention to the joint importance of positive and negative constructs in relation to psychological problems. For example, some cognitive-behavioural theories suggest that positive and negative cognitions about behaviour (cognitions about the advantages and disadvantages of a particular behaviour) have a central and causal role in the development and maintenance of binge eating and overeating (Beck, 2007; Cooper, Wells & Todd, 2004) and problematic substance and alcohol use (Beck, Wright, Newman & Liese, 1993).

Ryff and Singer (2003) emphasized the need to appreciate the dialectics between positive and negative aspects of living. They wrote that "human well-being is fundamentally about the joining of these two realms" (p. 279). Likewise, Snyder and Lopez (2007) commented that "future psychologists must develop an inclusive approach that examines both the weaknesses and the strengths of people" (p. 9). The joint important of positive and negative constructs has also begun to be recognised within the positive psychology movement, which initially solely focused on "positive" constructs. Examples include Peterson's (2006a) syllabus for his course on "positive psychology interventions," and Sheldon, Kashdan, and Steger's (2011) textbook *Designing the future of positive psychology: Taking stock and moving forward*.

A full discussion of all possible conceptualisations of the relationship between positive and negative constructs as they relate to psychological problems is beyond the scope of this thesis; the present discussion focuses on PCP's conceptualisation of the relationship between positive and negative constructs.

1.3. Positive and Negative Terminology and Conceptualisation

Before discussing positive and negative constructs any further, it is necessary to clarify what these terms mean in the context of this thesis. Throughout this thesis, "positive" and "negative" are used as they are used in the personality and social

psychology literature: to refer to the *valence* of particular variables. Most commonly, the terms "positive" and "negative" are used to refer to affective states. "Positive" is often, though not always, synonymous with adaptiveness, advantages, desirability, pleasantness, and "good"-ness. "Negative" is often, though not always, synonymous with maladaptiveness, disadvantages, undesirability, unpleasantness, and "bad"-ness.

PCP uses the terms "positive" and "negative" to denote variables that particularly characterise well-being and which have been explicitly focused on by positive psychology ("positive constructs") versus variables that particularly characterise psychological problems and which have been explicitly focused on by mental health researchers ("negative constructs").

Positive psychology and PCP both make the distinction between "positive" and "negative" as part of arguing for their novelty. As will be seen, positive psychology advocated for increasing awareness and research on "positive" constructs, whereas PCP advocated the value of integrating research on "positive" and "negative" constructs in order to make a call for the emerging positive psychology literature to be integrated with the voluminous existing literatures on mental health problems.

Some might assume that "positive" and "negative" are, respectively, synonymous with approach and avoidance motivations. Perspectives differ on this issue (as discussed in Chapter 3) and whether this is the case probably depends on the construct in question. For example, one dominant theory of self-regulation and emotion, which has much empirical support, suggests both positive and negative affect are *both* linked with approach and (separately) avoidance tendencies (e.g., Carver, 2003, 2004; Carver & Scheier, 1998).

As will be seen, sometimes "positive" and "negative" terminology are used to refer not just to separate research literatures but also to the valence of *opposite poles of the same variable*. In this context, "positive" and "negative" are used to refer to opposite poles of a single variable. Chapters 2 and 3 renalayse two commonly used clinical scales that contain a mixture of "negative" or "negatively-worded" items, which assess the presence of a negative construct (e.g., "I felt sad;" "I thought my life had been a failure") and "positive" or "positively-worded" items, which assess the presence of a positive construct (e.g., "I felt happy;" "I enjoyed life") in order to explore the hypothesis that certain psychological constructs form bipolar continua that each have a "positive" and a "negative" pole. More specifically, Chapter 2 tests whether depression forms one end of a depression-well-being bipolar continuum, where depression can be understood as the "negative" pole of this construct, and well-being as the "positive" pole of the same construct. Chapter 3 tests whether anxiety problems form one end of a bipolar anxiety-calmness continuum, where

anxiety can be understood as the "negative" pole of this construct, and calmness as the "positive" pole of the same construct.

1.4. Perceptions of Positive and Negative Constructs

The preceding discussion illustrates that the language of "positive" and "negative" often convey a value judgement. For instance, many people perceive that positive emotions such as excitement, enthusiasm, happiness, and calmness are good and to be sought out, whereas negative emotions such as sadness, guilt, and anxiety are bad and undesirable and to be avoided. Such perceptions underpin experiential avoidance – the tendency to avoid negative internal experiences – a key concept in several conceptualizations of psychological problems and theories of psychotherapy (Hayes, Strosahl & Wilson, 1999; Gamez et al., 2011). In fact, a large literature supports the notion that emotions of any kind are not inherently good or bad; evidencing that positive and negative emotions are equally functional as they both provide important information that can be used to guide behaviour and respond effectively to the environment (see Baumeister, Vohs, DeWall & Zhang, 2007).

These issues get at the heart of what this thesis is about, though, as it is individual *perceptions* that matter. This issue is particularly apparent in Chapters 4 and 5, which examine perceptions of SA and NSSI. Like any self-destructive behaviour, to an outsider, such behaviours may seem wholly undesirable – but to the person contemplating them or engaging in such behaviours, SA and NSSI make sense and confer important *perceived* advantages, for that person in their context, as Chapters 4 and 5 explore. Furthermore, although SA and NSSI may be viewed as conferring particular idiosyncratic advantages to the person who engages in these behaviours, to clinicians, friends, family, work colleagues, and society at large, NSSI is generally viewed as unconstructive at best and highly destructive and to be punished and stigmatised at worst (NICE, 2011). Thus, perceptions of what is "positive" versus "negative" may differ across informants.

Some scholars have argued that the language of "positive" and "negative" is therefore arbitrary and dependent on one's perspective and context (Held, 2016; Johnson & Wood, 2017; McNulty & Fincham, 2012; Wong, 2011). For example, optimism, the extent to which people hold generalized favourable expectancies for their future, might be considered a "positive" variable in that high levels of optimism have been linked to lots of advantages (Carver, Scheier & Segerstrom, 2010). However, optimism can be understood as a bipolar variable that resides on a continuum with pessimism and there appear to be at least some cases in which optimism has drawbacks (Carver et al., 2010).

Another example is empathy, the process of understanding another person's emotional state (Ferguson, 2016). Ferguson (2016) provides an interesting analysis of the adaptiveness of empathy and discusses how empathy may not be indiscriminately desirable. He notes that psychopaths have high levels of some kinds of empathy but are notably deficient in others. High levels of empathy are also sometimes associated with caring so much for other people that an individual neglects their own needs (Baron-Cohen, 2012), which in psychological therapy terms might be framed as "subjugating" one's needs and emotions (Young, Klosko & Weishaar, 2003). Such findings highlight the likelihood that no trait is universally good and adaptive for everyone all of the time and that the adaptiveness of particular traits is probably most accurately judged at the individual level in relation to each person's particular context.

1.5. Definitions of Positive and Negative Constructs in this Thesis

This discussion has highlighted some of the complexities involved in describing "positive" and "negative" psychological constructs. The language of "positive" and "negative," whilst imperfect, is parsimonious and relatively clear, and not using these terms has the potential to make the following discussion confusing as it would deviate from that used in the positive psychology and PCP literatures. Using this language is arbitrary, but doing so is necessary for clarifying what is the opposite of what. Throughout this thesis, "positive" will be used to refer to variables that particularly characterise well-being and which have been explicitly focused on by positive psychology. "Negative" will be used to refer to variables that particularly characterise psychological problems and which have been explicitly focused on by mental health researchers. As stated, when variables are bipolar, "positive" and "negative" are used to refer to opposite poles of the same variable, and Chapters 2 and 3 renalayse two commonly used clinical scales that contain a mixture of "negative" or "negatively-worded" items, which assess the presence of a negative construct (e.g., "I felt sad;" "I thought my life had been a failure") and "positive" or "positively-worded" items, which assess the presence of a positive construct (e.g., "I felt happy;" "I enjoyed life") in order to explore the hypothesis that certain psychological constructs form bipolar continua that each have a "positive" and a "negative" pole.

The discussion now turns to positive psychology and PCP in order to set the context for a discussion of whether positive and negative constructs are equally important for understanding the phenomenology of psychological problems.

1.6. Positive Psychology

Positive psychology has been defined as "the scientific study of positive experiences and positive individual traits, and the institutions that facilitate their development" (Duckworth, Steen & Seligman, 2005, p. 630). A full discussion of the emergence of the positive psychology movement is provided elsewhere (e.g., Linley, Joseph, Harrington & Wood, 2006; Seligman & Csikszentmihalyi, 2000; Johnson & Wood, 2017; Wood & Johnson, 2016). Briefly, positive psychology developed in the late 1990s, spearheaded by Martin Seligman. The rationale for its emergence was an argument that since World War II, mental health researchers had focused predominantly on the negative: on the assessment, understanding, and treatment of mental health problems, and critically, that this had occurred to the detriment of focusing on promoting well-being (the positive). Positive psychology aimed to raise awareness of the importance of researching "positive" traits: psychological traits and constructs that particularly characterise well-being (Gable & Haidt, 2005; Seligman & Csikszentmihalyi, 2000). It did this by taking an extreme position of *exclusively* focusing on positive traits and particularly aimed to help people who were functioning relatively well psychologically further enhance their well-being and performance.

1.6.1. Conceptual relationship between positive and negative

Positive psychology authors argued that the presence of positive traits is not simply the absence of negative traits, and that relieving negatives traits does not automatically bring about positive traits (Duckworth et al., 2005): "If the positive were just the absence of the negative, we would not need a positive psychology, just a psychology of relieving negative states" (p. 634). Thus, the variables studied by positive psychology researchers are seen to be separate to those studied by mental health researchers. That is, so-called "positive" and "negative" variables are each viewed as unipolar, ranging from zero to high levels of the particular construct in question.

1.6.2. Critique

Positive psychology and mental health researchers both aim to enhance well-being. Positive psychology researchers aim to investigate variables that support flourishing and well-being whilst mental health researchers aim to reduce distress (Johnson & Wood, 2017). The positive psychology movement undoubtedly experienced considerable success in making the study of "positive" variables interesting and important (Linley et al., 2006). However, over time, positive psychology has attracted criticism (e.g., see Joseph & Wood, 2010; Wood & Tarrier, 2010; Wood & Johnson, 2016), principally that it casts only a narrow spotlight on human experience by exclusively focusing on the positive (Joseph & Wood, 2010). Furthermore, the claim that psychology had focused exclusively on the

negative seemed to be rather overstated and dependent on one's theoretical and epistemological position. For example, this claim appears to overlook the contribution that the personality and social psychology literature has made to explaining human nature, which has framed with far less reference to "positive" versus "negative" constructs.

Additionally, plenty of existing, long-established clinical psychology theories and interventions include a focus on promoting "positive" constructs in order to ameliorate psychological problems. For example, a core aspect of Beck's cognitive therapy for depression (Beck, Rush, Shaw & Emery, 1979) involves helping people increase their engagement in pleasurable and fulfilling activities. Thus, as Mazzuchelli, Kane and Rees (2010) observe, there is little to distinguish behavioural activation from positive psychology interventions, other than the intention with which it was developed. Innumerable other psychological problem treatment models include a balance of relieving suffering (reducing the negative) and promoting well-being (promoting the positive), as discussed in Chapter 6. Some authors have also observed that positive psychology research has often been of relatively poor-quality and that the claims made have often exceeded the data (see Wood & Johnson, 2016). For example, several "positive interventions" have been developed and implemented that were not justified by basic science research (see Johnson & Wood, 2017; Wood & Johnson, 2016; Wood et al., 2010).

1.7. Positive Clinical Psychology

The criticisms levelled at positive psychology meant that, increasingly, scholars, even Seligman himself, began to move away from the either/or position that positive psychology had taken regarding positive and negative constructs to more of a position of integration. Just five years after the hallmark positive psychology paper (Seligman & Csikszentmihalyi, 2000), Seligman and others called for "a complimentary relationship between positive psychology and clinical psychology "as usual" (Duckworth et al., 2005, p.630) and several authors argued for the value of integrating positive psychology research with mental health research (e.g., Aspinwall & Staudinger, 2003; Larsen, Hemenover, Norris, & Cacioppo, 2003; Ryff & Singer, 2003). Over time, a zeitgeist shift has begun to occur towards a broader position that jointly focuses on alleviating psychological problems and promoting well-being (e.g., Ivtzan et al., 2015; Johnson & Wood, 2017; Joseph & Wood, 2010; McNulty & Fincham, 2012; Peterson, 2006b; Wood & Tarrier, 2010; Wood & Johnson, 2016), which some have labelled the "second wave" of positive psychology (Ivtzan et al., 2015; Wong, 2011).

One prominent component of this shifting zeitgeist was the inception of PCP, which, like others, called for positive psychology research to be integrated with the

voluminous evidence base concerned with understanding and treating psychological problems (Johnson & Wood, 2017; Wood & Tarrier, 2010; Wood & Johnson, 2016). The central thesis of PCP is that positive and negative psychological constructs are *equally important* for explaining psychological problems (Johnson & Wood, 2017; Joseph & Wood, 2010; Wood & Johnson, 2016; Wood & Tarrier, 2010). PCP argued that integrating positive psychology and mental health research would avoid reinventing the wheel by duplicating what is already known (Wood & Johnson, 2016), and would confer incremental prediction (Johnson & Wood, 2017; Joseph & Wood, 2010; Tarrier & Wood, 2010). For example, the positive psychology movement has exclusively focused on topics such as happiness, and much of the research has started from scratch on the assumption that there is no relevant existing research to draw upon (Joseph & Wood, 2010). However, this approach disregards existing research that is relevant to understanding "positive" variables (variables studied by positive psychology) that exists within the mental health field (Joseph & Wood, 2010; Johnson & Wood, 2017; Wood & Johnson, 2016) and other literatures.

More specifically, PCP predicts that depressive symptoms constitute one pole of depression-well-being continuum, where depression can be understood as the "negative" pole of this continuum, and well-being as the opposite, "positive" pole of the same continuum. If this prediction is correct, then the research base on well-being has a bearing on that for depressive symptoms, and vice-versa, because depression and well-being are the same construct, viewed through different lenses (Joseph & Wood, 2010)[1]. Chapters 2 and 3 discuss these issues in depth.

1.7.1. Conceptual relationship between positive and negative

The central argument of PCP is that the understanding and treatment of psychological problems needs to be based on researching, understanding, treating and fostering a balanced and equally weighted focus on positive and negative constructs (Wood & Tarrier, 2010). PCP puts forward two major predictions. The first is that *most psychological constructs are bipolar, consisting of a positive pole and a negative pole.* For example, tenacity is predicted to be bipolar with giving-up tendencies, flexibility with rigidity, self-acceptance with self-rejection, autonomy with subjugation, purpose in life with meaninglessness, environmental mastery with defeat, personal growth with psychological stagnation, optimism with pessimism, and gratitude with ingratitude (Johnson & Wood, 2017). In Chapter 3, a conceptual and methodological argument is put

[1] PCP originally proposed a depression-happiness continuum (Joseph & Wood, 2010; Tarrier & Wood, 2010) but has since moved away from this terminology due to debate regarding the meaning of "happiness"

forward that anxiety symptoms form one end of an anxiety-calmness continuum, where anxiety symptoms can be understood as the "negative" pole of this continuum, and calmness experiences as the opposite, "positive" pole of the same continuum (Joseph & Wood, 2010). The anxiety-calmness continuum is hypothesised to range from high anxiety symptoms, through a hypothetical zero point, to high calmness symptoms.

PCP's second major prediction is that a focus on variables that particularly characterise well-being ("positive" constructs) must equally complement a focus on variables that particularly characterise psychological problems ("negative" constructs) in mental health research and therapy because *positive constructs provide important unshared explanatory power in relation to psychological problems.* Specifically, positive constructs:

(1). Can predict psychological problems above and beyond negative characteristics because variables studied by positive psychology (e.g., gratitude) are different to those studied by mental health researchers (e.g., depression) and therefore share less variance;

(2). Buffer the impact of negative life events on distress, potentially preventing the development of psychological problems;

(3). Can be promoted in nonclinical populations to promote resilience;

(4). Can be fostered to treat psychological problems (interventions that aim to increase movement towards a "positive" pole of a particular construct can be used encourage movement away from the "negative" pole of the same construct, either in isolation or alongside traditional clinical interventions) (Johnson & Wood, 2017; Wood & Johnson, 2016; Wood & Tarrier, 2010).

1.7.2. Critique

PCP, like positive psychology, capitalised on an emerging zeitgeist. Neither research agenda is totally new. As stated, several authors prior to the inception of PCP advocated the value of integrating research from the positive psychology literature with existing research on mental health problems (e.g., Aspinwall & Staudinger, 2003; Larsen, Hemenover, Norris, & Cacioppo, 2003; Ryff & Singer, 2003). However, positive psychology and PCP both appear to have added to and probably extended what other authors have said by raising awareness of the importance of variables that might particularly explain well-being as they relate to psychological problems. PCP particularly focused on advocating the importance and value of integration, a message that is likely to provide benefits to the positive psychology and mental health field, as well as other fields.

Like positive psychology research, PCP is subject to several important criticisms. One important concern is that PCP has not always been explicit or consistent in discussing and framing its core predictions (Held, 2016), which has rendered its key messages

confusing at times. PCP has stated that "most characteristics have both positive and negative polls" (and are therefore bipolar) (Wood & Tarrier, 2010; p. 825) and that the "great majority of human functioning factors can be regarded as constituting a spectrum from negative to positive" (Johnson & Wood, 2017; p. 336). However, at other times, PCP has stated that positive and negative psychological constructs are *always* bipolar, for example implicitly with statements such as "characteristics range from low to high" (Wood & Johnson, 2016; p. 8), or explicitly with statements such as "all positive or negative factors have an inverse" (Johnson & Wood, 2017; p. 335) and "all factors fall on a continuum from positive to negative" (Johnson & Wood, 2017; p. 336). Overall, although PCP has been somewhat inconsistent in its phrasing, it generally seems to predict that most psychological constructs are bipolar, consisting of a positive pole and a negative pole. This thesis attempts to make an important contribution to the literature by clarifying and testing PCP's core messages and making them more explicit.

Another concern that can be raised is that there is very little direct evidence to support (or refute) PCP's central hypothesis that most psychological constructs are bipolar, having a positive pole and a negative pole. The authors place a large amount of importance (Joseph & Wood, 2010; Wood & Tarrier, 2010) on two studies they conducted (Joseph & Lewis, 1998; Wood, Taylor & Joseph, 2010), which demonstrated that depression can be understood as forming a continuum with well-being. They also put significant weight in a theoretical idea offered by Joseph and Wood (2010), who predicted that anxiety forms a continuum with calmness in order to support the notion that most psychological constructs are bipolar. However, no evidence was provided to support this hypothesised anxiety-calmness continuum. Chapters 2 and 3 of this thesis address the need for further direct evidence in support of PCP's first key prediction. These chapters test whether depression can be understood as forming a depression-well-being continuum and whether anxiety can be understood as forming an anxiety-calmness continuum.

Importantly, PCP has not made a specific prediction regarding the (linear or nonlinear) nature or form of the relationship between hypothesised bipolar continua and other psychiatric variables, making just an indirect remark that "it is possible that different therapies would be needed even if depression/happiness and anxiety/relaxation were continuums, if there was a non-linear relationship where different therapies worked better at different places on the continuum" (Joseph & Wood, 2010; p. 834). Chapters 2 and 3 of this thesis explain the importance of these analyses in detail and provide the first direct tests of this issue.

Chapters 2 and 3 examine the form of the relationship between the depression-well-being and other psychiatric variables (Chapter 2) and the calmness-anxiety continuum and other psychiatric variables (Chapter 3). Examining the form of these relationships, known as "phenomenological continuity" (Flett et al., 1997), would parallel work on stress, for example, which has established an inverted-U-shaped relationship between stress and memory function in that memory performance is impaired under conditions above or below optimal stress levels (Broadbent, 1965; Salehi, Cordero & Sandi, 2010; Yerkes & Dodson, 1908). Chapters 2 and 3 argue that these analyses have potentially significant implications for the conceptualization of psychological problems and for clinical practice.

As discussed, PCP's second major prediction states that variables that particularly characterise well-being ("positive constructs") provide important unshared explanatory power in relation to psychological problems (Johnson & Wood, 2017; Joseph & Wood, 2010; Wood & Johnson, 2016; Wood & Tarrier, 2010). This prediction requires some explanation and is potentially open to misinterpretation. It is fairly intuitive to see that when positive and negative constructs are separate (unipolar), positive constructs may explain incremental validity or even interact with negative constructs. However, this prediction is somewhat clear less when bipolar constructs are involved (Held, 2016), which have a positive pole and a negative pole. When this is the case, this prediction rests on the understanding that positive psychology and mental health researchers have tended to focus on *different constructs* (respectively "positive" and "negative" constructs) (Johnson & Wood, 2017). When a construct is bipolar (e.g., a depression-well-being continuum), each opposite pole (e.g., depression, well-being) cannot provide unshared explanatory power or add incremental validity in relation to the opposite pole *because the two constructs are the same thing.*

Different bipolar constructs are, however, potentially able to provide important unshared explanatory power in relation to psychological problems *because the constructs are different*. For example, gratitude may buffer against depression (see Wood, Froh & Geraghty, 2010), even though both constructs are hypothesised to be bipolar, because gratitude and depression are separate variables (Johnson & Wood, 2017). Gratitude has been predominantly focused on by positive psychology researchers and is hypothesised to be the "positive" pole of a gratitude/ingratitude bipolar continuum (Johnson & Wood, 2017). Depression has been predominantly focused on by mental health researchers and is hypothesised to be the "negative" pole of a depression-well-being bipolar continuum (Joseph & Wood, 2010; Wood et al., 2010).

Gratitude may buffer against depression, not because gratitude is a "positive" variable that has generally been studied by positive psychology researchers or because depression is a "negative" variable that has generally been studied by mental health researchers, but because the constructs are separate. The available measures of gratitude and depression happen to tap opposite ends of each bipolar continuum as a result of the separate epistemological position of the developers. That is, measures of gratitude just measure the "positive" pole of a potential gratitude/ingratitude continuum and, with a few exceptions, measures of depression generally just measure the "negative" pole of a potential depression-well-being continuum. If gratitude and depression, for example, are both found to be bipolar constructs (there has yet to be a test of a gratitude/ingratitude continuum), the effect of gratitude buffering against depression would need to be understood in terms of scoring high (at the positive end) on a gratitude continuum and scoring low (at the negative end) on a separate depression continuum.

PCP has been very clear about the importance and value of integrating positive psychology research ("positive" constructs) with mental health research ("negative" constructs) but the subtleties of its second core prediction, that positive and negative constructs each provide important unshared explanatory power in relation to psychological problems, may not have always come across to readers. It is hoped that the present analyses and discussion extends the existing explanations in this respect.

1.8. Positive and Negative Constructs and this Thesis

This thesis shines a spotlight on the largely untested central assumptions of PCP. Chapters 2 and 3 test: (1) whether depression and anxiety problems can each be understood as forming one pole of (separate) bipolar continua (respectively, a depression-well-being continuum and an anxiety-calmness continuum), (2) the (linear or nonlinear) form of the relationship between these hypothesised continua and other psychiatric variables. Chapters 4 and 5 test (3) whether some psychological variables are unipolar, and (4) whether positive psychological constructs explain additional variance over and above negative psychological constructs when predicting psychological problems (Johnson & Wood, 2016; Joseph & Wood, 2010; Wood & Johnson, 2016; Wood & Tarrier, 2010). The conceptual relationship between "positive" and "negative" constructs is therefore examined in relation to four psychological problems: depressive symptoms, anxiety-problems, SA and NSSI. Throughout this thesis, it is argued that testing these issues potentially has implications for how affective states are understood and how psychological problems are understood and treated.

1.8.1. Chapters 2 and 3

Chapters 2 and 3 examine whether positive and negative traits can, at times, be conceptualised as opposite ends of the same bipolar continuum using the examples of depressive symptoms (Chapter 2) and anxiety problems (Chapter 3). From the perspective of the literature on affect, one popular view suggests that positive and negative affect are independent dimensions that each range from zero to intense, such that, for example, anxiety or sadness are either present or absent (e.g., Cacioppo, Gardner & Berntson, 1999; Gray, 1990; Watson & Tellegen, 1999; Watson, Wiese, Valdya & Tellegen, 1999). Another popular view suggests that positive and negative affect are actually opposite poles of the same bipolar construct (e.g., Barrett & Russell, 1998; Carver & Scheier, 1998; Higgins, 1997; Joseph & Wood, 2010; Russell & Carroll, 1999a, b; Watson & Tellegen, 1999; Yik, Russell & Steiger, 2011). From the perspective of the literature on psychological problems, depressive and anxiety symptoms are usually conceptualised in research and practice as separate dimensions (Endler, Cox, Parker & Bagby, 1992; Feldman, 1993), but not as dimensions that are continuous with well-being and calmness (Joseph & Wood, 2010).

Chapter 2 explores whether depressive symptoms form one end of a depression-well-being bipolar continuum and Chapter 3 explores whether anxiety problems form one end of a bipolar anxiety-calmness continuum. As discussed below, Chapter 2 replicates a key study (Wood et al., 2010) that was used to argue for the notion that positive and negative constructs can be bipolar. Although this study was conducted in 2010, to the author's knowledge there has not been another direct test of this issue. However, this important finding (Wood et al., 2010) has been taken as key support in relation to PCP's argument for bipolarity, so a replication is much-needed. Chapter 3 provides the first test of an anxiety-calmness continuum, as hypothesised by Joseph and Wood (2010). Chapters 2 and 3 are important as they test whether the two most common psychological problems form (separate) bipolar continua.

The studies presented in these chapters also significantly extend previous findings by making the clinical importance of this topic more explicit. These studies test the (linear or nonlinear) form of the relationship between the depression-well-being and calmness-anxiety and continuums and other psychiatric variables. This form is critical to understanding anxiety problems, as discontinuities in relationships with other psychological conditions could be used to define a natural boundary of problematic depressive or anxiety symptoms. Chapters 2 and 3 potentially provide important implications for how we understand affect and how we conceptualize psychological problems and deliver mental health services, particularly whether it would be *equally*

important to reduce negative constructs (depressive and anxiety symptoms) and promote positive constructs (calmness and well-being).

1.8.2. Chapters 4 and 5

Chapters 4 and 5 examine whether positive and negative constructs can, at times, be conceptualised as separate, using the example of self-injurious cognitions (SICs). SICs are cognitions specifically about engaging in self-injurious behaviour (SIB) such as "People think that my suicide attempt(s) are selfish" or "Nonsuicidal self-injury helps me escape negative emotions." "SIB" refers to behaviours to deliberately physically injure oneself.

Chapters 4 and 5 describe the development and validation of the Suicide Attempt Beliefs Scale (SABS) and the Nonsuicidal Self-Injury Beliefs Scale (NSIBS), two new self-report measures of SICs. The SABS was developed to measure beliefs about suicide attempts (SA). The NSIBS was developed to measure beliefs about engaging in nonsuicidal self-injury (NSSI). Chapters 4 and 5 also explore whether SA and NSSI are *each* characterised by positive and negative SICs. "Positive SICs" are the idiosyncratic individual and interpersonal perceived advantages of SIB (e.g., "My problems are so serious that SA is the only option" [individual]; "NSSI makes people take my problems seriously" [interpersonal]). "Negative SICs" are the idiosyncratic individual and interpersonal perceived disadvantages of SIB (e.g., "NSSI makes my problems worse" [individual]; "People judge and criticise my SA" [interpersonal]).

Previous measures of SICs confined their focus to measuring particular aspects of positive SICs. Thus, negative SICs were previously outside of the measurement spotlight and a comprehensive measure of positive and negative SICs was not available. Structural analyses will explore the conceptual relationship between positive and negative SICs.

1.9.The Replication Crisis

Although this thesis is conceptually underpinned by PCP a methodological theme that is apparent in every study/chapter, is that of replication. Given the centrality of this methodological theme to this thesis, and to psychological science more generally, what follows is an overview of the key tenets of replicability and the "replication crisis", followed by a discussion of how each thesis chapter incorporates replication (in order to make explicit how replication features as a key methodological narrative for this thesis). It is important for contemporary PhD theses to demonstrate an awareness of the replication crisis because this so-called "crisis" is a major topical debate currently facing all of psychological science. Moreover, Everett and Earp (2015) have argued that the PhD thesis is the ideal opportunity to contribute replication attempts to the field.

1.9.1. What is the Replication Crisis?

Simply speaking, the "replication crisis" literature has drawn attention to the fact that some key scientific findings do not replicate (e.g., Ioannidis, 2005; Open Science Collaboration, 2012), which some have taken as evidence that there may be a "replication crisis." This literature explicitly emphasises and encourages replication and takes a critical and thoughtful look at how scientific replication can be defined.

The replication crisis literature has particularly focused on findings in psychological science and biomedicine, although the concept of replication has obvious relevance to all of science because reproducibility – the extent to which consistent results are observed when individual studies are repeated – is of course one of the defining features of science. In fact, some authors contend that replication is "the cornerstone of science" because only replication can adjudicate whether a single study reporting an original result represents a true finding or a false positive result (Simons, 2014, p. 76).

Replication matters and it is often assumed in science. The worst case scenario for the replication crisis debate is that the "scientific literature is too good to be true" (Bakker, van Dijk & Wicherts, 2012, p. 543); one cannot believe anything that is published, and psychology is filled with mostly invalid and false explanations (Braver et al., 2014). If particular key findings do not replicate, researchers may have fruitlessly expended resources in the pursuit of false leads, and policy and practice may have been based on false information (Open Science Collaboration, 2012). A low estimate of the reproducibility of current findings potentially brings the quality of published research into question, motivating further investigations of reproducibility, and catalysing changes in practice and publishing standards (Bertamini & Munafò, 2012; LeBel & Peters, 2011). On the flip side, devoting resources to confirmation instead of innovation is a poor investment if the original findings are valid (Open Science Collaboration, 2012); hence why replication has been de-emphasised in terms of grant funding and publishability. A high estimate of the reproducibility of current findings could therefore nullify concerns of a replication "crisis" by bolstering confidence in conventional research and peer-review practices (Open Science Collaboration, 2012).

1.9.2. Types of Replication

Broadly speaking, there are two types of replication designs. "Exact" or "direct" replications are replications that operationalize both the independent and the dependent variable(s) in exactly the same way as the original study. The purpose of such replication attempts is to verify the original findings (Stroebe & Strack, 2014). Exact replications are both useful and necessary in applied research. For example, repeatedly applying exactly

the same procedure is critical for establishing the efficiency of a specific intervention, where, depending on the subject matter (e.g., cancer treatment), a lack of reliability could have fatal consequences (Stroebe & Strack, 2014).

"Conceptual" replications aim to replicate findings by testing hypotheses that were examined in a particular study using a different study design (Stroebe & Strack, 2014). The purpose of conceptual replications is to see whether a given phenomenon is apparent across a range of conditions (Schmidt, 2009). Although conceptual replications can be very informative (especially as content for narrative reviews), they may not be able to identify false positive results because differences between original and replication studies may be due to procedural differences (Maxwell, Lau & Howard, 2015). For this reason, the replication crisis literature has emphasised exact or direct replications, especially of experiments.

1.9.3. The Zeitgeist of Replication

Although there is general agreement that replicability is important, there are relatively few successful replication studies (Open Science Collaboration, 2012), which indicates either that findings are false (i.e., those replication studies that have been attempted fail to produce supportive findings) or that journals are biased against replication studies whatever their result (Giner-Sorolla, 2012; Neuliep & Crandall, 1990, 1993; Rosenthal, 1979).

Everett and Earp (2015) and Earp (2015) have argued that the replication crisis can be understood as a disciplinary "social dilemma," wherein the collective interests of the field and society are at odds with the private interests of individual researchers. This social dilemma arises because of the contradiction that replication attempts are viewed as important indications of rigorous and reliable science by the field and society, whilst at the same time the individuals who actually conduct those replications "are looked down on as bricklayers and not [as] advancing [scientific] knowledge" (Makel et al., 2012, p. 537). Thus, conducting replication attempts is seen as professionally costly for individuals because such endeavours involve shifting energy and resources away from projects that involve original thinking and new ideas onto projects that are less likely to be published (Earp, 2015; Everett & Earp, 2015). This problem is argued to be particularly acute for early career academics who experience the "publish or perish" pressure more acutely than established academics (Everett & Earp, 2015).

1.9.4. Replication Attempts

The "replication crisis" as it relates to psychology was most famously recently illustrated in the Reproducibility Project, which involved 270 researchers attempting to

directly replicate almost 100 psychology studies (Open Science Collaboration, 2012, 2015). Each team selected a key finding from a single study for replication. Researchers solicited feedback on their research design from the original authors before collecting data in order to identify potential barriers to replication. Where threats to replication were identified, efforts were made to either remedy these or else such factors were coded as potential predictors of reproducibility and written into the replication report (Open Science Collaboration, 2012, 2015). The outcome is that according to one of several measures of reproducibility, only 36% of results could be confirmed; by another statistical measure, 47% could. The results of the Reproducibility Project therefore suggested a major problem with reproducibility, potentially bringing psychological science into disrepute.

Another large-scale illustration of the "replication crisis" was provided via a recent survey of 1576 researchers that was conducted by the journal Nature. This survey revealed that >70% have tried and failed to reproduce another scientist's experiments, and >50% have failed to reproduce their own experiments (Baker, 2016). However, <31% of respondents believed that a failure to reproduce published results means that the result is probably wrong, and most believed that they still trust the published literature (Baker, 2016).

1.9.5. Reproducibility is Complex

The initial furore of a replication "crisis" is increasingly being accompanied by a more thoughtful consideration of the complexities of replication, with increasing discussions of the nuances and facets of reproducibility. For example, the American Psychological Association (2015) emphasized the possibility that hidden moderators rendered the Reproducibility Project's replications ineffective, and an independent re-examination of the Reproducibility Project's results found that when correcting for sampling error, power, and bias, the data in fact indicate that the reproducibility of psychological science is quite high (Gilbert, King, Pettigrew & Wilson, 2016; for a reply, see Anderson et al., 2016).

The issue of reproducibility is, in fact, complex and a direct replication is unlikely to produce a perfect replication (Patil, Peng & Leek, 2016). Several authors have also highlighted that there are no established criteria or consensus for deciding that a finding has indeed replicated or what replication actually means (Gómez, Juzgado, & Vegas, 2010; Valentine et al., 2011). "Successful" replication might be defined narrowly (i.e., obtaining the same statistically significant effect as the original study) or broadly (i.e., obtaining a directionally similar, but not necessarily statistically significant, result) (Open Science

Collaboration, 2012), or with no reference to the statistical significance of findings (discussed below).

A failure to replicate a finding does not conclusively indicate that the original finding was false because there are myriad possible reasons for findings not replicating. These include: Type II error (some true findings will fail to replicate purely by chance, with an occurrence rate of $1-\beta$); insufficient power (the actual size of the effect is lower than originally reported, making it more difficult to detect); "researcher degrees of freedom" (researcher decision-making when collecting and analysing data, e.g., Should more data be collected? Which conditions should be compared?); publication bias; fraud; problems with the design, implementation, or analysis of the original or replication study; a failure to recognize and document the circumstances and social context in which research takes place; changes in the population over time; and other known and unknown factors (Baker, 2016; Braver et al., 2014; Cesario, 2014; Cohen, 1962; Cumming, 2014; Earp & Trafimow, 2015; Etz & Vandekerckhove, 2016; John, Loewenstein & Prelec, 2012; Klein et al., 2012; Maxwell et al., 2015; Meehl, 1990; Open Science Collaboration, 2012; Patil et al., 2016; Rosenthal, 1979; Simmons, Nelson & Simonsohn, 2011; Stroebe & Strack, 2014).

Some authors have also argued that direct replications arguably contribute little to scientific knowledge because they simply demonstrate that the original outcome was reproducible (Stroebe & Strack, 2014). The researcher may not be any wiser as to whether the original study was a good test of the theory to be tested because even though the experiment may have been poorly designed, a faithful replication might result in the same finding (Stroebe & Strack, 2014). However, if a particular hypothesis is supported using different designs (a conceptual replication), the researcher will have gained additional information, bolstering their trust in an underlying theory (Stroebe & Strack, 2014). Additionally, failures to replicate a research finding cannot be used as indicators of fraud, nor can successful replications be invoked as indication that the original study was honestly conducted (Stroebe & Strack, 2014). For example, Diederik Stapel, a prominent social psychology researcher who committed major scientific fraud, wrote in his autobiography how he was always pleased when his invented findings were replicated (Stroebe & Strack, 2014).

1.9.6. Robust Methods for Establishing Replication

Some solutions have been proposed to the complexities of replication, including: (i) Conducting multiple, independent, adequately powered replication studies (Maxwell et al., 2015); (ii) using Bayesian statistics to compare the weight of evidence for competing (null

and alternative) hypotheses (Braver et al., 2014; Etz & Vandekerckhove, 2016); (iii) using meta-analysis to combine the results of replication studies (Braver et al., 2014); (iv) developing theories to explain inconsistent findings (Stroebe & Strack, 2014); and (v) directly testing whether findings are actually heterogeneous or due to publication bias and explaining substantive heterogeneity in terms of moderator variables using meta-analysis (Borenstein et al., 2009; Lipsey & Wilson, 2001).

It is worth noting the importance of power in replication attempts (point (i) in the above list), as this is a critical issue where replication is concerned, and one that is fairly easily remedied at the design stage of research. Several of the discussions of the issue of replicability have highlighted the fact that replication attempts have been underpowered (e.g., Braver et al., 2014; Maxwell et al., 2015; Patil et al., 2016). This has sometimes occurred, for instance in the Reproducibility Project, because power estimates were based on observed effect sizes in the original studies (Etz & Vandekerckhove, 2016). This means that, in the event that the original effect sizes were inflated, the sample size recommendations from prospective power analysis will be underestimates, and therefore replication studies will tend to find mostly weak evidence as well (Etz & Vandekerckhove, 2016; Maxwell et al., 2015).

The notion that research is often underpowered has been long lamented. Cohen (1962) estimated that the typical power of published psychological studies does not exceed 0.5. More recently, for example, Sedlmeier and Gigerenzer (1989) found a similar value among studies conducted in the subsequent 25 years, indicating that the low power of most earlier published behavioural research was not rectified by Cohen's findings. If all psychological research were carried out with power near .95 or even .80, replication problems would arguably diminish. True effects would virtually always achieve significance and virtually always be replicable, and researchers would start being more interested in effect sizes (Braver et al., 2014).

A re-analysis of the findings of the Reproducibility Project concluded that the apparent failure to replicate many target effects can be adequately explained by overestimation of effect sizes (or overestimation of evidence against the null hypothesis) due to small sample sizes and publication bias in the psychological literature (Etz & Vandekerckhove, 2016). The authors used Bayesian analysis and found that the majority of replication attempts (64%) did not provide strong evidence for either the null or the alternative hypothesis in either the original or the replication, and no replication attempts provided strong evidence in favour of the null. In all cases where the original paper provided strong evidence but the replication did not (15%), the sample size in the

replication was smaller than the original. Where the replication provided strong evidence but the original did not (10%), the replication sample size was larger (Etz & Vandekerckhove, 2016).

Elaborating on the use of meta-analysis to examine the replicability of findings (points (iii) and (v) in the above list), the meta-analytic perspective very usefully discourages dichotomous decisions regarding successful replication based on $p < .05$, and instead advocates continuous descriptive criteria such as effect size estimates and their confidence intervals (Braver et al 2014; Cumming, 2014). As is well-known, the .05 criterion for what constitutes convincing, reliable, "statistically significant" evidence is completely arbitrary. In Rosnow and Rosenthal's (1989) words, "surely, God loves the .06 nearly as much as the .05" (p. 1277). Additionally, "Fisher offered the idea of *p* values as a means of protecting researchers from declaring truth based on patterns in noise. In an ironic twist, *p* values are now often manipulated to lend credence to noisy claims based on small samples" (Gelman & Loken, 2014, p. 460).

Applying meta-analysis to the issue of replication, our confidence should increase that the phenomenon under study is genuine and true (replicable) to the degree that the combined criterion remains significant, whereas indices of heterogeneity (e.g., Q and I^2) remain small and nonsignificant. Larger Q and I^2 values should prompt the search for plausible moderator variables which explain heterogeneity among replication attempts (Borenstein et al., 2014; Braver et al., 2014; Lipsey & Wilson, 2001). Multiple independent replications of the same study are therefore needed to definitively evaluate replication (Patil et al., 2016). One large-scale replication project called the "Many Labs" Replication Project (Klein et al., 2014) did just this, replicating findings 35 or 36 times (which would guarantee adequate power to identify true effects) and then pooling the results. This approach contrasts with that adopted in the Reproducibility Project mentioned earlier and indeed produced a much more heartening result whereby 85% of the original studies were successfully replicated.

1.10. Replication and This Thesis

1.10.1. Chapter 2

Chapter 2 presents a direct replication of a study by Wood et al. (2010), who used the Center for Epidemiologic Studies-Depression Scale (CES-D; Radloff, 1977) to re-examine the structure of depressive symptoms. Wood et al. (2010) used confirmatory factor analysis (CFA) to control for positively and negatively worded items (method effects) in an adult and an older adult sample and found that the CES-D measures a bipolar continuum that ranges from well-being to depression (Wood et al., 2010). Chapter 2

provides a direct replication of these analyses and findings in a large, diverse sample of adolescents (*N* = 4,138).

1.10.2. Chapter 3

Chapter 3 presents a direct replication of a study by Vautier and Pohl (2009), who used the State-Trait Anxiety Inventory (STAI; Spielberger et al., 1970) to re-examine the structure of state and trait anxiety. The authors used structural equation modelling (SEM) to control for item wording (method effects) in the French adaptation of the STAI. Trait and state STAI scales were each found to measure one construct (Vautier & Pohl, 2009). Chapter 3 discusses how this study is limited because it utilised an exclusively adult sample who completed the French adaptation of the STAI. It was therefore unclear whether these results would generalize beyond adults and to the original and more commonly used English language version of the scale. Chapter 3 attempts to directly replicate the analyses and findings of Vautier and Pohl (2009) using two large samples (*N* = 4,138 and 1,824).

1.10.3. Chapter 4

Chapter 4 describes the development of the SABS and NSIBS, two new measures of beliefs about SA and beliefs about NSSI. Replication is a critical facet of scale development and validation and I explore whether important findings replicate in several ways. In the development paper (Chapter 4), the SABS and NSIBS are developed across six large, independent samples of people with lived experience of SIB (total *N* = 3,313). A series of separate Exploratory Factor Analyses (EFA) and Confirmatory Factor Analyses (EFA) will be conducted in order to locate a robust and generalizable factor structure for the SABS and NSIBS.

1.10.4. Chapter 5

The validation paper for the SABS and NSIBS (Chapter 5) explicates the internal consistency, test retest reliability, convergent, discriminant, criterion, and incremental validity of the SABS and NSIBS. Replication also features strongly here, in that all of the psychometric analyses will be conducted in relation to at least two independent samples. Conclusions will then be drawn based on the consensus of results across different samples, which again, provides some confidence that results are relatively robust and generalizable.

1.11. Published Material

As mentioned, this thesis is structured around four core chapters (Chapters 2-5), with an Introduction (Chapter 1) and General Discussion (Chapter 6). Each chapter has been written and presented in a format that is appropriate for publication in a high ranking, peer-reviewed journal. Chapters 2 and 3 present manuscripts that have been published in the *Journal of Personality and Social Psychology* and the *Journal of Affective Disorders*,

respectively. Chapters 4 and 5 present manuscripts that will be submitted to *Psychological Assessment*. This thesis format and structure was specifically chosen in order to explicitly demonstrate how completing this PhD led to the development of a wide range of research and academic skills and knowledge. (The learning that occurred as a result of completing this PhD is discussed in more depth in Chapter 6).

1.11.1. Note on Collaboration and Published Material

As the thesis chapters present manuscripts that have either been published or which will shortly be submitted to academic peer-reviewed journals, a note on collaboration and published material is therefore necessary.

The research presented in this thesis was completed under the supervision of several individuals. The vast majority of the guidance and input, and my development as a researcher, came from my principal supervisor, Prof Alex M. Wood. Prof Alex M. Wood and Dr Peter J. Taylor provided information, advice, guidance, and feedback on conceptualisation, writing, and statistical analyses for Chapters 2 and 3, and are therefore recognised as co-authors. Profs Alex M. Wood, Ronan E. O'Carroll, and Rory C. O'Connor provided information, advice, guidance, and feedback on conceptualisation, writing, and statistical analyses for Chapters 4 and 5, and are therefore recognised as co-authors.

The data collection that took place for Chapters 4 and 5 was conducted entirely by me. All statistical analyses were undertaken solely by me. Each manuscript was written almost entirely by me, with occasional editing by my supervisors during the very final stages of manuscript preparation.

CHAPTER 2

2. The Centre for Epidemiologic Studies-Depression (CES-D) Scale Measures a Continuum From Well-Being to Depression: Testing Two Key Predictions of Positive Clinical Psychology

2.1.Abstract

Two core but untested predictions of Positive Clinical Psychology (PCP) are that (1) Many psychiatric problems can be understood as one end of bipolar continua with well-being, and (2) that reducing psychiatric symptoms will provide an equal (near linear) decrease in risk for several other psychiatric variables, irrespective of position on continua. We test these predictions in relation to a purported well-being/depression continuum, as measured by the Center for Epidemiologic Studies-Depression (CES-D), a popular measure of depressive experiences in research and clinical practice. A large (N = 4,138), diverse sample completed the CES-D, which contains a mixture of negatively worded and positively worded items (e.g., "I felt sad," "I enjoyed life"). The latter are conventionally reverse scored to compute a total score. We first examined whether purportedly separate well-being and depression CES-D factors can be reconceptualised as a bipolar well-being/depression continuum. We then characterised the (linear or nonlinear) form of the relationship between this continuum and other psychiatric variables. Both predictions were supported. When controlling for shared method bias amongst positively worded items, a single factor well-being/depression continuum underlies the CES-D. Baseline levels on this continuum are found to have near linear relationships with changes in anxiety symptoms, aggression, and substance misuse over time, demonstrating that moving from depression to well-being on the CES-D provides an equal decrease in risk for several other psychological problems irrespective of position on the continuum. The CES-D does not measure well-being as comprehensively as established scales of well-being. Results support calls for mental health services to jointly focus on increasing well-being and reducing distress, and point to the value of early intervention and instilling resilience in order to prevent people moving away from high levels of well-being.

This chapter has been published as: Siddaway, A. P., Wood, A. M., & Taylor P. J. (2017). The Centre for Epidemiologic Studies-Depression (CES-D) scale measures a continuum from well-being to depression: Testing two key predictions of Positive Clinical Psychology. *Journal of Affective Disorders, 213,* 180-186.

2.2.Introduction

Well-being is becoming an increasingly central focus of international policy (e.g., Department of Health, 2009; Mental Health Commission of Canada, 2009). The positive psychology movement was proposed in order to raise awareness of the importance of researching psychological traits and constructs that promote well-being (Gable & Haidt, 2005; Seligman & Csikszentmihalyi, 2000). This literature has burgeoned and recent years have seen an increasing shift towards a broader position that jointly focuses on alleviating psychological problems *and* promoting well-being (e.g., Ivtzan et al., 2015; Joseph & Wood, 2010; McNulty & Fincham, 2012; Peterson, 2006b; Wood & Tarrier, 2010), which some have labeled the "second wave" of positive psychology (Ivtzan et al., 2015; Wong, 2011).

One prominent component of this shifting zeitgeist has been the inception of Positive Clinical Psychology (PCP), which has called for positive psychology research to be integrated with the voluminous evidence base concerned with understanding and treating psychological problems (Wood & Tarrier, 2010, as clarified in Johnson & Wood, in press). Numerous articles (e.g., Johnson & Wood, in press; Lomas, 2015; Joseph & Wood, 2010; Wood & Tarrier, 2010; Wong, 2011) and at least three books (Ivtzan et al., 2015; Peterson, 2006b; Wood & Johnson, 2016) have now summarized evidence which supports this integration. These demonstrate, for example, that: (i) constructs studied by positive psychology researchers can independently predict psychological problems over and above clinical constructs; (ii) that constructs studied in positive psychology can confer resilience to psychological problems; and (iii) that interventions which aim to move people towards well-being can also be used to help people move away from psychological problems, either in isolation, or alongside clinical interventions.

Although this progress is promising, the notion that positive psychology research can and should be integrated with the existing evidence base for psychological problems remains largely untested. This study explores this issue by testing two core predictions made by PCP using the example of depressive experiences. The first prediction to be tested here is the idea that many psychological problems in fact form continua with well-being. A well-being/psychological problem continuum would indicate that research on either area has implications for the opposite pole (Joseph & Wood, 2010; Wood & Tarrier, 2010) and would suggest that the language of "positive" and "negative" is arbitrary and dependent on one's perspective and context (Johnson & Wood, 2017; McNulty & Fincham, 2012; Wood & Johnson, 2016). This conceptualization would align with the substantial research base demonstrating that psychological problems are best-viewed as continuous rather than

discrete categories (e.g., see Bentall, 2003; Haslam et al., 2012; Markon et al., 2011) and which has examined continuously distributed transdiagnostic constructs and mechanisms (e.g., see Harvey et al., 2004).

We test this key prediction in relation to a hypothesized depression/well-being continuum using the Center for Epidemiologic Studies-Depression (CES-D) scale (Radloff, 1977). The CES-D is one of the most frequently used self-report measures of depressive experiences (Santor et al., 2006) and there is extensive support for its psychometric properties (Ensel, 1986; Radloff, 1977; Roberts, 1980; Shean & Baldwin, 2008). The fact that the CES-D contains a mixture of negatively worded items (e.g., "I felt sad;" "I thought my life had been a failure") and positively worded items (e.g., "I felt happy;" "I enjoyed life") led to the proposal that it could be re-conceptualised as a depression/well-being continuum (Joseph, 2006, 2007). It was argued that for a score of zero to occur on the CES-D, a person would have to give all of the negatively worded items (e.g., "I felt sad") the lowest possible score ("rarely of none of the time") *and* all of the positively worded items (e.g., "I enjoyed life") the highest possible score ("most or all of the time"). For such a person it would be misleading to state that they have merely indicated an absence of depressive experiences; such an individual has also clearly indicated the *presence* of well-being (Joseph, 2006, 2007; Joseph & Wood, 2010).

One existing study has tested whether the CES-D can be re-conceptualized as a well-being/depression continuum (Wood et al., 2010). The authors found that when account for item wording using structural equation modeling, the CES-D can indeed be understood as measuring a bipolar continuum that ranges from well-being to depression. The authors established these findings using separate adult and older adult samples and presented evidence that the well-being items (e.g., "I felt happy;" "I enjoyed life") demonstrate convergent validity with the well-validated Scales of Psychological Well-being (Ryff, 1989). Given the potential practical importance of the suggestion that depression forms a bipolar continuum with well-being, and the increasing emphasis on replicating scientific findings to ensure that they are robust and generalizable, our first aim was to replicate the structural analyses of this previous study using a large, diverse sample of adolescents.

The second prediction made by PCP that we test here is the idea that moving along the well-being/depression continuum (total score on the CES-D) towards well-being will provide an equal decrease in risk for several other psychological problems, irrespective of position on the continuum (Wood & Johnson, in press; Wood & Tarrier, 2010). We test this prediction by examining the form of the relationship between the depression/well-

being continuum and other psychological problems over time, which Flett et al. (1997) referred to as "phenomenological continuity." That is, continuity in the relationship between psychological problems and their antecedents, concomitants, or sequalae. Accordingly, even if depression is relatively continuous in a psychometric sense, its relationship with associated variables could be relatively discontinuous or nonlinear in form, defining a natural boundary of depressive experiences (Markon, 2010).

The benefit of these analyses is that they make the clinical importance of this topic more apparent and explicit than merely examining whether a well-being/depression continuum exists. One possibility, for example, is that there is no relationship between the well-being/depression continuum and other psychological problems up to a particular point (e.g., throughout the range of the well-being pole), after which the detrimental consequences of depression begin to manifest. Evidence of this relationship would corroborate the current emphasis in mental health services on alleviating and treating psychological problems. This conceptualization of depression (and other mental health) problems underpins psychiatric nomenclature and, as a result, psychiatric and psychological interventions tend to be stopped at the point of problem absence.

An alternative possibility is a linear relationship between the well-being/depression continuum and other psychological problems throughout the range of the continuum. This relationship would be apparent if depressive experiences increase at a constant rate along with other psychological problems, without any threshold defining a change in association. Evidence of this relationship would simultaneously highlight the importance of treating depression (because as depressive experiences increase, so do other psychological problems) and emphasize the usefulness of fostering well-being (because as well-being increases, psychological problems decrease). Such evidence would be consistent with calls from professional bodies (e.g., The British Psychological Society, 2010) and the mental health recovery movement (Anderson, Oades & Caputi, 2003) for mental health services to focus not just on tackling psychological problems but also on fostering well-being and helping people live a valued, meaningful life.

We examine these two key predictions of PCP using a large population-based archival dataset, which, by implication, involved variability in the latent entity, thereby minimizing the likelihood of systematic sampling bias, which could have been introduced had we used a purely community or clinical sample (Waller & Meehl, 1998). For example, using an undergraduate sample could introduce a systematic sampling bias since only those individuals functioning well enough to attend classes would be studied (Hankin et al., 2005). Likewise, focusing entirely on clinical individuals may limit variability in

depressive experiences (Hankin et al., 2005) as clinical samples often exhibit more severe symptoms and greater comorbidities than population-based samples (Newman et al., 1998).

2.3.Method

2.3.1. Participants

The sample comprised 4,138 adolescents and adults aged 13-21 years from Hawai‘i. These individuals took part in the five-year longitudinal Hawaiian High Schools Health Survey (HHSHS) study conducted by the National Center on Indigenous Hawaiian Behavioral Health (NCIHBH). This sample provides a broad spread of ages, ethnicities, socioeconomic status,' and gender (Andrade et al., 2006; Hishinuma et al., 2000). Participants for the HHSHS study were sampled from five high schools which were selected from both urban and rural areas to obtain a representative sample of adolescents residing in Hawaii. Students who provided assent completed the survey in their classrooms under the supervision of their teachers. Parents of students younger than 18 years old were notified of the study by mail and given an opportunity to refuse participation. Data collected during the 1992/1993 (N = 4,164), 1993/1994 (N = 4,182), and 1994/1995 (N = 1,433) school years were used in this study. There was some missing demographic information and incomplete questionnaire responses (see Andrade et al., 2006; Hishinuma et al., 2001) and we multiply imputed missing data as best practice (discussed below).

2.3.2. Measures

Center for Epidemiological Studies-Depression (CES-D; Radloff, 1977). CES-D responses capture the frequency of feelings and behaviours over the past seven days and are rated on a 4-point scale ranging from 0 (rarely or none of the time) to 3 (most or all of the time). The CES-D contains 20 items that are summed so that scores have a potential range from 0 to 60, with higher scores indicating greater frequency of depressive experiences (Radloff, 1977). Numerous studies have examined the factor structure of the CES-D (McArdle et al., 2001; Prescott et al., 1998; Radloff, 1977) and these have generally suggested that negative and positive items load onto separate factors (see Shafer, 2006, for a review). There is extensive support for the psychometric properties of the CES-D, including high internal consistency in community and clinical populations (Cronbach's αs .85 - .90; Ensel, 1986; Radloff, 1977; Roberts, 1980); convergent validity with other popular measures of depressive experiences such as the Patient Health Questionnaire-9 (r = .85; Amtmann et al., 2014) and Beck Depression Inventory-II (Shean & Baldwin, 2008); and divergent validity from, for example, aggression (r = .44) and substance use (r = .24; Makini et al., 1996). A cutoff score of 16 has been found to have sensitivity and specificity

rates of 86.7 and 76.6 for identifying depressed individuals, whereas a cutoff score of 21 has a sensitivity and specificity rate of 73.0 and 96.1 (Shean & Baldwin, 2008). The CES-D demonstrated excellent internal consistency (α = .89).

CES-D items representing depression and well-being were generally positively skewed.

Distribution of distributions of average responses to the STAI State items

	Number of days per week symptoms experienced, on average			
Items	<1 day	1-2 days	3-4 days	5-7 days
Well-being	1306	1182	821	550
Depression	2180	939	483	275

State-Trait Anxiety Inventory-Trait scale (STAI; Spielberger et al., 1970). Trait anxiety is seen as a relatively stable individual difference in the tendency to respond to situations perceived as threatening with elevation in state anxiety (Spielberger et al., 1970). Items are rated on a 4-point frequency scale based on "how you generally feel." Support for the psychometric properties of the STAI has been extensive (see Spielberger, 1989; Spielberger et al., 1970). The trait scale demonstrates excellent internal consistency (average Cronbach's αs > .89); excellent test–retest reliability at multiple time intervals (average *r* = .88; Barnes et al., 2002); adequate convergent validity with other measures of anxiety (Spielberger, 1989); and divergent validity from, for example, aggression (*r* = .38) and substance use (*r* = .19; Knight et al., 1983). The STAI demonstrated excellent internal consistency (α = 89).

Substance Abuse Subtle Screening Inventory—Adolescent version (SASSI-A; Miller, 1990). The SASSI-A is a brief screen for substance use, impairment, and dependency arising from substance use. The SASSI-A has been shown to have good psychometric properties, including acceptable internal consistency in the current sample (α = .74); and divergent validity from anxiety (*r* = .19), depression (*r* = .24), and aggression (*r* = .33; Makini et al., 1996). It has also been shown to concord with a diagnosis of substance abuse and dependency on the Diagnostic Interview Schedule for Children (Nishimura et al., 2001), predict counselor DSM-III diagnoses for dually diagnosed adolescent inpatients (Piazza, 1996), and predict adolescent chemical dependency (Risberg et al., 1995). The SASSI-A demonstrated good internal consistency (baseline/1992/1993 school year: α = .88; Time 1/1993/1994 school year: α = .88; Time 2/1994/1995 school year: α = .86).

Braver Aggressiveness Dimension Scale (BADS; Braver et al., 1986). The BADS is a 14-item abbreviated self-report measure of child and adolescent aggression. It was derived from the longer Youth Self-Report scales (YSR: Achenbach, 1991), the self-

report version of the Child Behaviour Checklist. Items selected for the BADS were those items from the YSR which were significantly more likely to be endorsed by clinically-diagnosed, conduct disordered children and adolescents. The BADS has good psychometric properties, including one year test-retest stability (r = .61) and divergent validity (Makini et al., 1996). It demonstrated good internal consistency (baseline/1992/1993 school year: α = .85; Time 1/1993/1994 school year: α = .70; Time 2/1994/1995 school year: α = .64).

Missing Data

There were substantial amounts of missing data in the Hawaiian dataset. 6.4% of all values were missing for the 1992/1993 school year, 52.2% of all values were missing for the 1993/1994 school year, and 22.53% of all values were missing for the 1994/1995 school year. The missingness was not completely at random (MCAR). We addressed this potential problem by multiply imputing missing data on all variables at the item level using SPSS version 21.0 (IBM Corp, 2012). Multiple imputation (MI) is increasingly advocated as the optimal approach for dealing with missing data (Graham, 2009; Schafer & Graham, 2002; Shrive et al., 2006). When MI has been compared with alternative methods of handling incomplete data (e.g., single imputation methods, complete-case analyses, maximum likelihood approaches), it has been shown to generate less biased estimates that have more statistical efficiency (e.g., Crawford et al, 1995; Donders et al., 2006; Liu & Gould, 2002; Tang et al., 2005). There is also evidence indicating that MI performs well across different circumstances, such as small samples, very large multiple regressions, and when there are large amounts of missing data (Graham & Schafer, 1999).

MI works by generating plausible missing values multiple times based on the distribution of the observed data. Random components are incorporated into these estimated values to reflect their uncertainty. This procedure creates a set of ‘‘complete’’ data sets with no missing values. Analyses are then run separately on each data set, and the results are pooled across datasets using multiple imputation combining rules (Enders, 2010; Graham, 2009). The purpose of MI is not to obtain the individual values themselves but to estimate unbiased parameter estimates of the data set as a whole (Graham, 2009). We followed recommendations to match the number of imputations to the fraction of missing information because progressively larger numbers of imputed datasets are needed to maximize power in subsequent significance testing (Bodner, 2008; Graham et al., 2007; White et al., 2011).

Statistical Analysis

Analyses were conducted using SPSS version 21.0 (IBM Corp, 2012) and R (R Development Core Team, 2009). The CFAs were performed using the R lavaan package, version 0.5-18 (Rosseel, 2012). Positively worded CES-D items were reverse scored for all analyses. Complete cases were used for these analyses and three CFA models were tested. Model 1 was the standard two factor model consisting of separate negatively worded (e.g., "I felt sad") and positively worded (e.g., "I enjoyed life") items, which were allowed to correlate. Model 2 was a single factor model with all items loading on a single factor. Model 3 also featured a single substantive well-being/depression factor, however the positively worded items were allowed to cross-load onto a second methodological artefact factor which accounted for additional residual inter-correlation between these items.

As CES-D data are ordinal, we employed WLSMV estimation (Flora & Curran, 2004). Acceptable fit was operationalized as Root Mean Squared Error of Approximation (RMSEA) $\leq$ .08, Comparative Fit Index (CFI) $\geq$ .90, and Tucker Lewis Index (TLI) $\geq$ .90. Good fit was operationalized as RMSEA $\leq$ .06, CFI $\geq$ .95, and TLI $\geq$ .95 (Hu & Bentler, 1999). The fit of competing CFA models were compared using: (i) Akaike's Information Criterion (AIC; calculated using Maximum Likelihood CFA), which tests the relative fit of competing models after adjusting for parsimony (lower AICs indicate less information loss and thus a superior model), and (ii) CFI, using a .002 cutoff (Meade et al., 2008). Both approaches are argued to be superior to using the chi-square statistic to compare model fit because this statistic is known to be highly sensitive to sample size (see Meade et al., 2008).

Hierarchical ordinary least squares (OLS) regressions were used to explore linear and nonlinear relationships between CES-D scores, treated as a single factor (all items summed to produce a total score), and outcome variables. Regressions were conducted using MI data. In each analysis, Step 1 involved fitting a model whereby CES-D total scores had a linear relationship with the outcome variable measured at the same time (1992/1993 school year), or measured at follow-up 1 or 2 years later (1993/1994, 1994/1995 school years), whilst controlling for scores on the outcome variable at baseline (hence it was the change in outcome that we were predicting). Steps 2 and 3 tested whether adding a nonlinear term (squared and cubed CES-D total scores) made a significant improvement to the amount of variance explained. Improvement in model fit was based on ΔR^2. Statistically significant deviations from linearity were graphed in order to visually display relationships, using unstandardized regression coefficients. This also clarified whether nonlinearity was substantive.

2.4. Results

2.4.1. Comparison of CFA Models

Table I shows that the two factor model (Model 1) demonstrated an improvement in fit over the one factor model (Model 2), replicating previous findings regarding the factor structure of the CES-D (see Shafer, 2006). Model 3 (a well-being/depression continuum) demonstrated an improvement in fit again, thereby replicating the findings of Wood et al. (2010). The AIC statistic, which accounts for model complexity, and the change in model fit according to CFI (Meade et al., 2008), both pointed to the superiority of Model 3, as hypothesized. These results indicate that when shared method bias amongst positively worded items is controlled for, a single factor underlies the CES-D items. Furthermore, as we argued in the Introduction, Model 3 was also favored on theoretical grounds because endorsing positively worded items, which are usually reverse-scored (e.g., "I felt happy," "I enjoyed life"), does not merely indicate the absence of depressive experiences; it actually indicates the *presence* of well-being (Joseph, 2006, 2007; Joseph & Wood, 2010).

2.4.2. Exploration of Linear and Nonlinear Relationships with Outcome Variables

A series of regression analyses were conducted to explore the form of the relationship between the CES-D and outcome variables over time (Table II). Three of the eight regression models showed statistically significant nonlinear relationships for Step 2. However, the squared term accounted for very little additional variation above and beyond the linear main effect (0.7%, 0.1%, 0.2%). One of the eight regression models showed statistically significant nonlinear relationships for Step 3. However, again, the cubed term accounted for very little additional variation (0.6%). Thus, in all cases the nonlinear term failed to make any substantive improvement to the original linear model and results provide only very weak evidence of a nonlinear relationship. The observed nonlinearity appears to be of statistical but not clinical significance and the relationship between the well-being/depression continuum and outcome variables appears to be near linear over time. Statistically significant nonlinear relationships were also explored graphically (Figure 1). The graphs reveal only subtle variation away from perfect linearity.

Given the large proportion of missing data, we conducted a robustness check of our regression analyses using complete cases (see Supplementary material for Chapter 2). These results were almost identical (in terms of ΔR^2 values), again finding that baseline levels on the well-being/depression continuum have a near linear relationship with changes in trait anxiety, aggression, and substance misuse over time (regression coefficients are

larger than those in Table II as multiple imputation estimates relationships more conservatively).

Overall, the regression analyses and graphs provide evidence that the well-being/depression continuum (CES-D total scores) has a near linear relationship with outcome variables over time. Any nonlinearity appears to be of statistical but not practical or clinical significance. These results support our prediction that moving from depression to well-being on the CES-D provides an equal decrease in risk for several other psychological problems, irrespective of position on the well-being/depression continuum.

Table I. Comparison of Three Mean- and Variance-Adjusted Weighted Least Squares CFA Models.

	Model fit					
Model	X^2	*df*	AIC	TLI	CFI	RMSEA
1. Two factor	4209.891*	169	163260.390	.920	.929	.082
2. Single factor	8025.236*	170	165187.856	.845	.861	.115
3. Single factor, method variance factor	3921.218*	166	163175.121	.924	.934	.080

Note. N = 4138; CES-D completed during 1993/1994 school year; analyses are reported to three decimal places for clarity; AIC statistic produced from Maximum Likelihood CFA; * $p < .001$.

Table II. Results of Regression Analyses Comparing Linear and Nonlinear Effects of CES-D upon Change in Outcome.

Step	Outcome variables	*B*	*SE B*	*β*	*ΔR²*
	Trait anxiety				
1992/1993 school year (*N* = 4,069)					
1	Constant	11.243	.194		
	Total CES-D score	.702	.010	.744***	.554***
2	Constant	9.347	.310		
	Total CES-D score	.952	.034	1.008***	
	Total CES-D score squared	-.006	.001	-.277***	.007***
3	Constant	6.815	.446		
	Total CES-D score	1.493	.077	1.581***	
	Total CES-D score squared	-.034	.004	-1.615***	
	Total CES-D score cubed	.000	.000	.813***	.006***
	Aggression				
1992/1993 school year (*N* = 4,069)					
1	Constant	1.784	.124		
	Total CES-D score	.249	.006	.523***	.274***
2	Constant	1.969	.199		
	Total CES-D score	.225	.022	.472***	
	Total CES-D score squared	.001	.000	.054	.000
3	Constant	1.695	.292		
	Total CES-D score	.283	.051	.595***	
	Total CES-D score squared	-.002	.002	-.234	
	Total CES-D score cubed	.000	.000	.175	.000
1993/1994 school year (*N* = 4,101)					
1	Constant	5.334	.145		
	1992/1993 BADS total score	.354	.018	.380***	
	Total CES-D score	.027	.008	.062***	.176***
2	Constant	4.994	.224		
	1992/1993 BADS total score	.354	.018	.381***	
	Total CES-D score	.072	.024	.162*	
	Total CES-D score squared	-.001	.001	-.105*	.001*
3	Constant	4.660	.314		

	1992/1993 BADS total score	.354	.018	.380***	
	Total CES-D score	.144	.053	.324**	
	Total CES-D score squared	-.005	.003	-.482	
	Total CES-D score cubed	.000	.000	.228	.001
1994/1995 school year (N = 4,101)					
1	Constant	7.897	.211		
	1992/1993 BADS total score	.244	.023	.302***	
	Total CES-D score	.020	.009	.051*	.115***
2	Constant	7.573	.254		
	1992/1993 BADS total score	.244	.023	.303***	
	Total CES-D score	.062	.023	.161*	
	Total CES-D score squared	-.001	.000	-.116*	.002*
3	Constant	7.366	.331		
	1992/1993 BADS total score	.244	.023	.302***	
	Total CES-D score	.106	.052	.277*	
	Total CES-D score squared	-.003	.002	-.384	
	Total CES-D score cubed	.000	.000	.163	.001
	Substance misuse				
1992/1993 school year (N = 4,069)					
1	Constant	.262	.038		
	Total CES-D score	.045	.002	.336***	.113***
2	Constant	.259	.062		
	Total CES-D score	.045	.007	.339***	
	Total CES-D score squared	.000	.000	-.003	.000
3	Constant	.284	.090		
	Total CES-D score	.040	.016	.299*	
	Total CES-D score squared	.000	.001	.253	
	Total CES-D score cubed	.000	.000	.151	.000
1993/1994 school year (N = 4,101)					
1	Constant	1.122	.041		
	1992/1993 SASSI-A total score	.397	.016	.436***	
	Total CES-D score	.006	.002	.050*	.208***
2	Constant	1.059	.059		
	1992/1993 SASSI-A total score	.397	.016	.436***	

	Total CES-D score	.014	.006	.119*	
	Total CES-D score squared	.000	.000	-.071	.001
3	Constant	1.009	.086		
	1992/1993 SASSI-A total score	.397	.016	.436***	
	Total CES-D score	.025	.015	.207	
	Total CES-D score squared	-.001	.001	-.277	
	Total CES-D score cubed	.000	.000	.125	.000
1994/1995 school year (N = 4,101)					
1	Constant	1.754	.056		
	1992/1993 SASSI-A total score	.221	.020	.297***	
	Total CES-D score	.005	.002	.052*	.098***
2	Constant	1.681	.071		
	1992/1993 SASSI-A total score	.220	.020	.296***	
	Total CES-D score	.015	.006	.149*	
	Total CES-D score squared	.000	.000	-.101	.001
3	Constant	1.666	.089		
	1992/1993 SASSI-A total score	.220	.020	.296***	
	Total CES-D score	.018	.014	.181	
	Total CES-D score squared	.000	.001	-.176	
	Total CES-D score cubed	.000	.000	.046	.000

Note: CES-D = Center for Epidemiological Studies-Depression; STAI = State Trait Anxiety Inventory; BADS = Braver Aggressiveness Dimension Scale; SASSI-A = Substance Abuse Subtle Screening Inventory—Adolescent version; CES-D scale completed during 1992/1993 school year; analyses are reported to three decimal places for clarity; *$p < .05$, ***$p < .001$.

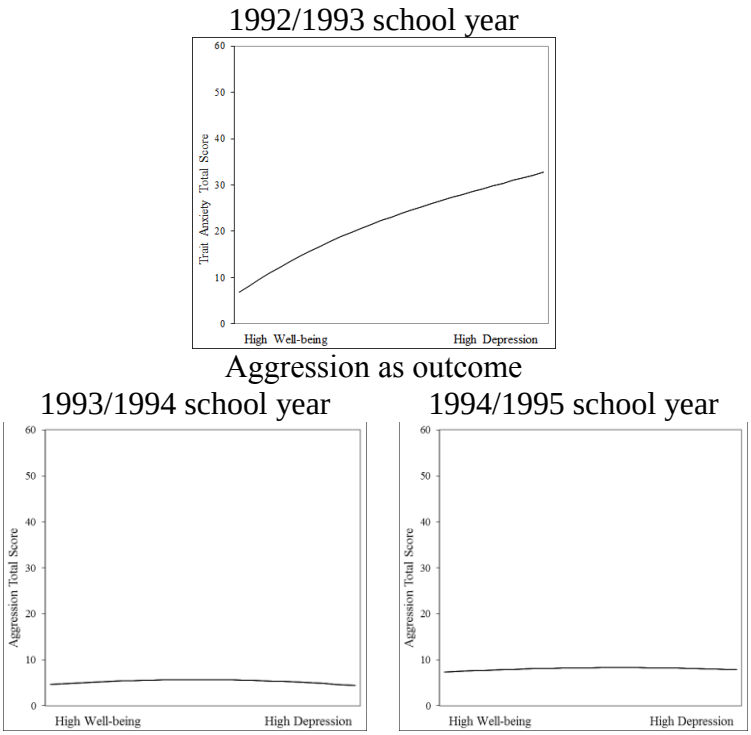


Figure 1. Line graphs plotting unstandardized nonlinear regression lines for statistically significant ΔR2 values. Total CES-D scores predict outcome variables at different time points; CES-D scale completed during 1992/1993 school year.

2.5. Discussion

This study tested two key predictions made by PCP using the example of depressive experiences. We first examined whether the CES-D can be reconceptualised as measuring a continuum that ranges from well-being to depression using a large, heterogeneous sample. These analyses replicated the novel findings of one previous study that used samples of adults and older adults (Wood et al., 2010) and extended them to adolescents. This replication was important because of the potential importance of the suggestion that depression forms a bipolar continuum with well-being. We used a large, ethnically diverse sample that is representative of adolescents residing in Hawaii in order to minimize the likelihood of systematic sampling bias, which could have been introduced had we used a purely community or clinical sample. Taken together with the previous findings (Wood et al., 2010), there is now evidence that the CES-D measures a well-being/depression continuum in adolescents, adults, and older adults. The CES-D total score provides an indication of depressive symptom severity that is based on a combination of the presence/absence of depressive experiences and the presence/absence of well-being. Importantly, this conceptualisation contrasts with other views regarding the inter-relationship between psychological problems and well-being, for instance challenging the Complete State Model of Mental Health's (Keyes, 2003, 2007) assertion that "mental illness" and well-being form separate continua.

The theoretical underpinning for our analyses was the suggestion that for a score of zero to occur on the CES-D, a person would have to give all of the negatively worded items (e.g., "I felt sad") the lowest possible score ("rarely of none of the time") and all of the positively worded items (e.g., "I enjoyed life") the highest possible score ("most or all of the time;" (Joseph, 2006, 2007; Joseph & Wood, 2010). Responding in this way would not, therefore, merely indicate an absence of depressive experiences; it would actually indicate the *presence* of well-being (Joseph, 2006, 2007; Joseph & Wood, 2010). Conceptualizing depression as one end of a well-being/depression continuum has the advantage of removing arbitrary divides between maladaptive and adaptive. This evidence could therefore be used to make an argument against mental health stigma, since everyone resides somewhere on the well-being/depression continuum. Our conceptualization of depression suggests that individuals who are currently experiencing depressive experiences are not qualitatively different from anyone else; at the moment they are simply further along the well-being/depression continuum towards high depressive experiences.

This is the first study to examine the form of the relationship between the well-being/depression continuum and other psychological problems over time. Our results

demonstrate that baseline levels on the well-being/depression continuum have a near linear relationship with outcome variables measured at the same time and one and two years later. It therefore seems desirable to score highly on the well-being pole of the continuum because the higher an individual's well-being, as measured by the CES-D, the less likely that individual is to experience several psychological problems (anxiety symptoms, aggression, problematic substance use) over time. These results point to the value of jointly focusing on increasing well-being and reducing distress, as doing so promotes resilience from developing other psychological problems. They also point to the importance of early intervention in order to prevent people moving away from high levels of well-being. As depression and well-being form a single continuum, fostering high levels of well-being means that individuals have further to move before the onset of depressive experiences; in effect raising the threshold for the onset of depressive experiences. These results join evidence that anxiety can be understood as residing on a calmness-anxiety continuum that also has near linear relationships with other psychological problems over time (Siddaway, Taylor & Wood, in press).

Until recent years, the dominant zeitgeist in mental health services has very understandably been to focus on the negative aspects of life and how these can be reduced and alleviated. Our results add to an emerging evidence-base which suggests that we may need to change the way we think about mental health problems. Our results support calls for mental health services to jointly focus on increasing well-being and reducing distress (e.g., The British Psychological Society, 2010) and the World Health Organisation's (WHO; 2004) conceptualization of mental and physical health as "a state of complete physical, mental and social well-being and not merely the absence of disease or infirmity." PCP's explicit call for a balanced and equal focus on the positive and negative aspect of life (Wood & Johnson, in press; Wood & Tarrier, 2010; Wood & Johnson, 2016) aligns well with the view offered by the mental health recovery movement, where "recovery" involves "the establishment of a fulfilling, meaningful life and a positive sense of identity" (Anderson et al., 2003).

People will of course continue to present to mental health services wanting to address their psychological problems. However, the present results could be used to inform an evidence-based discussion regarding when to stop interventions and the advantages and disadvantages of doing so. The evidence presented here suggests that it may be beneficial for clinical interventions to continue beyond the mere absence of depressive symptoms. Although this is clearly less cost-effective in the short-term, it could prove to be the most cost-effective option long-term.

The current results need to be interpreted in light of several limitations, many of which stem from using an archival dataset. Further research is clearly needed to consolidate arguments regarding a more balanced and equally weighted focus on treating problems and increasing well-being. One important limitation of our results arises because the CES-D was not designed to measure a well-being/depression continuum; therefore the well-being items do not measure this pole of the continuum as comprehensively as established scales of well-being. Indeed, the CES-D contains unequal numbers of negatively worded and positively worded items, disproportionately weighting the depression end of the continuum. Developing and validating a scale that specifically measures the well-being/depression continuum is therefore expected to provide additional predictive validity over the CES-D.

It will also be important to establish the form of the relationship between the well-being/depression continuum and psychological problems other than those measured here, as well as measures of positive psychological functioning. Although we used a large, heterogeneous sample, it is possible that the nature and form of the well-being/depression continuum may change over time or vary across particular groups, for instance as a result of cultural changes in how mental health problems are understood. Determining whether and how conceptualisations and inter-relationships of psychological problems and well-being change over time represents a core area of ongoing research. An economic cost value analysis that tests the implications of the well-being/depression continuum conceptualization and when it is most cost effective to stop interventions is particularly needed.

2.6. Supplementary Material for Chapter 2

Table S1. Results of Regression Analyses Based on Complete Cases Comparing Linear and Nonlinear Effects of CES-D upon Change in Outcome.

Step	Variables	*B*	*SE B*	*β*	ΔR^2
	Trait anxiety as outcome				
1992/1993 school year (*N* = 3,759)					
1	Constant	10.645	.199		
	Total CES-D score	.725	.011	.747***	.558***
2	Constant	8.797	.316		
	Total CES-D score	.975	.035	1.005***	
	Total CES-D score squared	-.006	.001	-.270***	.007***
3	Constant	6.858	.449		
	Total CES-D score	1.403	.079	1.445***	
	Total CES-D score squared	-.028	.004	-1.316***	
	Total CES-D score cubed	.000	.000	.643***	.004***
	Aggression as outcome				
1992/1993 school year (*N* = 3,910)					
1	Constant	1.862	.121		
	Total CES-D score	.250	.006	.527***	.278***
2	Constant	1.999	.194		
	Total CES-D score	.232	.021	.488***	
	Total CES-D score squared	.000	.000	.041	.000
3	Constant	1.845	.277		
	Total CES-D score	.266	.049	.560***	
	Total CES-D score squared	-.001	.002	-.129	
	Total CES-D score cubed	.000	.000	.105	.000
1993/1994 school year (*N* = 1,935)					
1	Constant	1.274	.152		
	1992/1993 BADS total score	.564	.020	.585***	
	Total CES-D score	.039	.010	.085***	.402***
2	Constant	1.162	.236		
	1992/1993 BADS total score	.564	.020	.585***	
	Total CES-D score	.054	.027	.119*	
	Total CES-D score squared	.000	.001	-.036	.000

3	Constant	.819	.334		
	1992/1993 BADS total score	.564	.020	.585***	
	Total CES-D score	.132	.060	.288*	
	Total CES-D score squared	-.005	.003	-.440	
	Total CES-D score cubed	.000	.000	.250	.001
1994/1995 school year (*N* = 925)					
1	Constant	1.714	.248		
	1992/1993 BADS total score	.542	.034	.521***	
	Total CES-D score	.030	.016	.060	.309***
2	Constant	1.309	.382		
	1992/1993 BADS total score	.540	.034	.519***	
	Total CES-D score	.088	.045	.178*	
	Total CES-D score squared	-.001	.001	-.124	.001
3	Constant	.950	.542		
	1992/1993 BADS total score	.540	.034	.518***	
	Total CES-D score	.171	.099	.345	
	Total CES-D score squared	-.006	.005	-.516	
	Total CES-D score cubed	.000	.000	.241	.001
	Substance misuse as outcome				
1992/1993 school year (*N* = 3,781)					
1	Constant	.223	.038		
	Total CES-D score	.044	.002	.334***	.111***
2	Constant	.221	.061		
	Total CES-D score	.044.	.007	.336***	
	Total CES-D score squared	-.000	.000	-.002	.000
3	Constant	.250	.088		
	Total CES-D score	.038	.015	.288*	
	Total CES-D score squared	.000	.001	.112	
	Total CES-D score cubed	.000	.000	-.070	.000
1993/1994 school year (*N* = 1,817)					
1	Constant	.228	.042		
	1992/1993 SASSI-A total score	.608	.021	.579***	
	Total CES-D score	.010	.002	.080***	.376***
2	Constant	.236	.068		

	1992/1993 SASSI-A total score	.608	.021	.579***	
	Total CES-D score	.009	.008	.071	
	Total CES-D score squared	.000	.000	.009	.000
3	Constant	.200	.097		
	1992/1993 SASSI-A total score	.608	.021	.580***	
	Total CES-D score	.017	.017	.138	
	Total CES-D score squared	.000	.001	-.150	
	Total CES-D score cubed	.000	.000	.098	.000
1994/1995 school year (N = 866)					
1	Constant	.454	.068		
	1992/1993 SASSI-A total score	.474	.037	.414***	
	Total CES-D score	.013	.004	.103***	.211***
2	Constant	.353	.107		
	1992/1993 SASSI-A total score	.474	.037	.414***	
	Total CES-D score	.027	.012	.216*	
	Total CES-D score squared	.000	.000	-.118	.000
3	Constant	.337	.154		
	1992/1993 SASSI-A total score	.474	.037	.414***	
	Total CES-D score	.031	.028	.244	
	Total CES-D score squared	-.001	.001	-.186	
	Total CES-D score cubed	.000	.000	.042	.000

Note: CES-D = Centre for Epidemiological Studies-Depression; STAI = State Trait Anxiety Inventory; BADS = Braver Aggressiveness Dimension Scale; SASSI-A = Substance Abuse Subtle Screening Inventory—Adolescent version; CES-D scale completed during 1992/1993 school year; *p < .050, ***p < .001.

Table S2. Table Summarizing Which Centre for Epidemiological Studies-Depression (CES-D) Items Measure Depression and Well-Being.

	Depression	Well-Being
Item 1	I was bothered by things that usually don't bother me	
Item 2	I did not feel like eating; my appetite was poor	
Item 3	I felt that I could not shake off the blues even with help from my family or friends	
Item 4		I felt that I was just as good as other people
Item 5	I had trouble keeping my mind on what I was doing	
Item 6	I felt depressed	
Item 7	I felt that everything I did was an effort	
Item 8		I felt hopeful about the future
Item 9	I thought my life had been a failure	
Item 10	I felt fearful	
Item 11	My sleep was restless	
Item 12		I was happy
Item 13	I talked less than usual	
Item 14	I felt lonely	
Item 15	People were unfriendly	
Item 16		I enjoyed life
Item 17	I had crying spells	
Item 18	I felt sad	
Item 19	I felt that people dislike me	
Item 20	I couldn't get going	

CHAPTER 3

3. Re-Conceptualizing Anxiety as a Continuum that Ranges from High Calmness to High Anxiety: The Joint Importance of Reducing Distress and Increasing Well-Being

3.1. Abstract

We first replicate a study by Vautier and Pohl (2009), who used the State-Trait Anxiety Inventory (STAI) to re-examine the structure of anxiety. Using two large samples (N = 4,138 and 1,824), we also find that state and trait anxiety measure continua that range from high calmness to high anxiety. We then significantly extend previous findings and make the clinical importance of this topic more explicit by characterizing the (linear or nonlinear) form of the relationship between the calmness-anxiety continuum and other psychiatric variables for the first time. This form is critical to understanding anxiety problems, as discontinuities in relationships with other psychological conditions could be used to define a natural boundary of problematic anxiety. Baseline levels on the calmness-anxiety continuum are found to have a near linear relationship with changes in depression, aggression, and substance misuse over time. Taken together, these results indicate the joint importance and usefulness of treating anxiety problems and promoting calmness, as doing so may promote resilience from developing other psychiatric conditions. Psychiatric and psychological interventions that are grounded in this continuum conceptualization would logically be stopped when an individual reports experiencing high levels of calmness. Our results point to the usefulness of early intervention and prevention (when people begin to move away from high calmness) and instilling resilience (by providing interventions to move people towards high calmness).

This chapter has been published as: Siddaway, A. P., Taylor, P. J., & Wood, A. M. (2017). Re-conceptualizing anxiety as a continuum that ranges from high calmness to high anxiety: The joint importance of reducing distress and increasing well-being. *Journal of Personality and Social Psychology*.

3.2. Introduction

The STAI is "the most widely used device to measure anxiety across cultures" (Lonner & Ibrahim, 1989: p. 317) and there is extensive support for its psychometric properties (see Spielberger, 1989; Spielberger & Diaz-Guerrero, 1986; Spielberger et al., 1970). Strong consensus over decades of research has supported a four factor structure that consists of state and trait "anxiety present" and "anxiety absent" factors (Bernstein & Eveland, 1982; Hishinuma, Miyamoto, Nishimura & Nahulu, 2000; Mook, Van der Ploeg & Kleijn, 1992; Spielberger, Vagg, Barker, Donham & Westberry, 1980; Suzuki, Tsukamoto & Abe, 2000; Vagg, Spielberger & O'Hearn, 1980; Vigneau & Cormier, 2008). "Anxiety absent" items (e.g., "I feel calm") are reverse scored. As currently defined, a total score on the state and trait scales therefore involves a combination of the "presence" and "absence" of anxiety symptoms.

Although there has been extensive support for a four factor structure to the STAI, there are both conceptual and statistical reasons to suggest that this issue needs to be revisited, with implications for how we understand the structure of anxiety. The conceptual argument was provided by Joseph and Wood (2010), who observed that reverse scored STAI items appear to assess calmness and relaxation (e.g., "I am cool, calm and collected" [trait anxiety]; "I feel calm" [state anxiety]), rather than merely the absence of anxiety problems. They pointed out that for the lowest possible score on the STAI to occur, a person would have to give all of the positively worded items (e.g., "I feel calm") the highest possible score ("Almost always") and all of the negatively worded items (e.g., "I feel anxious") the lowest possible score ("Almost never"). On this basis, they suggested that the lowest score on the STAI does not just indicate the absence of anxiety problems, as has been convention in research and practice for decades; it actually indicates the presence of calmness and relaxation. The STAI, as conventionally-coded, could therefore be (re)conceptualised as ranging from high calmness to high anxiety. Although this interpretation of the STAI is new, the idea that anxiety might form a continuum with calmness has previously been recognized by the circumplex model of affect (e.g., Kuppens et al 2013; Russell, 1980; Russell & Carroll, 1999a, b; Yik et al., 2011) and by Carver and Scheier's (1998) model of affect[2].

The statistical argument for re-examining the factor structure of the STAI is as follows. It is common practice for self-report psychiatric scales to include some positively

2 Carver and Scheier (1998) suggest that the avoidance system underlies this continuum and that calmness or anxiety are experienced depending on an individual's perceived effectiveness (or ineffectiveness) in moving away from threat.

worded items that are reverse scored in order to compute a total score (Woods, 2006). The rationale is that these items measure the same construct as the negatively worded items (psychiatric symptoms), whilst reducing the tendency for respondents to agree more than disagree (acquiescence bias), or respond according to their general feeling about the topic rather than the specific content of the items (a response set; Green et al., 1993; Woods, 2006). However, there is evidence that including positively worded items in psychiatric scales can inadvertently lead to the existence of a separate "method factor" that is not substantively meaningful (Green et al., 1993; Vautier & Pohl, 2009; Woods, 2006). Two Monte Carlo studies demonstrate that the existence of two factors would be inferred from the normal methods of exploratory and confirmatory factor analysis if only 10% of respondents respond carelessly to positively worded items (Schmitt & Stults, 1985; Woods, 2006). When separate factors within a scale respectively contain only positively worded and negatively worded items, the potential substantial importance of each factor is confounded with potential artifactual effects (Woods, 2006). These observations have led to a growing consensus amongst methodologists that factor analytic models need to account for item wording when demonstrating the existence of separate substantive factors in scales that include reverse scored items (Woods, 2006).

One study to date has examined the factor structure of the STAI whilst accounting for the potential influence of positively worded items (Vautier & Pohl, 2009). The authors used structural equation modelling (SEM) to control for item wording (method effects) in the French adaptation of the STAI. Trait and state STAI scales were each found to measure one construct rather than separate "anxiety present" and "anxiety absent" factors (Vautier & Pohl, 2009).

Although these structural findings have important implications for the scoring and interpretation of the STAI, the broad applicability and clinical significance of these findings was limited in at least two important respects. First, this study utilised an exclusively adult sample who completed the French adaptation of the STAI. It is therefore unclear at present whether these results would generalize beyond adults and to the original and more commonly used English language version of the scale.

Second, the form of the relationship between the calmness-anxiety continuum and other psychiatric variables remains to be established. This form is critical to understanding the boundaries of anxiety problems, as it may be that anxiety has a stronger relationship with other difficulties when it reaches a particular level, thereby demarcating what is psychopathological versus what is not (Markon, 2010). Flett, Vredenburg and Krames (1997) referred to this as "phenomenological continuity," that is, continuity in the

relationship between psychopathology and its antecedents, concomitants, or sequalae. Thus, even if anxiety is relatively continuous in a psychometric sense, its relationship with associated variables might be relatively discontinuous or nonlinear in form, thereby defining a natural boundary of problematic anxiety (Markon, 2010).

The Current Study

Given the potential practical importance of understanding the structure of anxiety, the present study sought to replicate and extend the findings of Vautier and Pohl (2009). We begin by using two large, diverse samples to test whether a hypothesized continuum is apparent in the most commonly used, English language version of the STAI when the influence of positively worded items is accounted for using SEM. We then extend the findings of Vautier and Pohl (2009) by characterizing the (linear or nonlinear) form of the relationship between the calmness-anxiety continuum and other psychiatric variables for the first time. This approach parallels work on stress, for example, which has established an inverted-U-shaped relationship between stress and memory function in that memory performance is impaired under conditions above or below optimal stress levels (Broadbent, 1965; Salehi, Cordero & Sandi, 2010; Yerkes & Dodson, 1908). These analyses make the clinical importance of this topic more apparent than merely examining whether a continuum exists. An analysis of related variables in several domains is required to establish phenomenological continuity (Flett et al., 1997); we use measures of depression, substance abuse, and aggression, as these are often comorbid with anxiety (Kendler, Prescott, Myers & Neale, 2003; Mineka, Watson & Clark, 1998). We predict that moving from anxiety to calmness will provide an equal (near linear) decrease in risk for several other psychiatric variables over time, irrespective of position on the calmness-anxiety continuum (total score on the STAI).

Clarifying the form of the relationship between the calmness-anxiety continuum and other psychiatric variables has potentially significant implications for the conceptualization of psychological problems and clinical practice. For example, it is possible that there is no relationship between the calmness-anxiety continuum and other psychiatric variables up to a particular point (e.g., throughout the range of the calmness pole), after which anxiety symptoms come to have an increasingly detrimental effect. Evidence in support of this conceptualisation would point to the well-known deleterious effects of anxiety symptoms and therefore support the current emphasis in mental health services on alleviating and treating anxiety problems. This conceptualization of anxiety (and other mental health) problems underpins psychiatric nomenclature and, as a result,

psychiatric and psychological interventions tend to be stopped at the point of anxiety problem absence.

An important alternative conceptualisation is one in which the relationship between the calmness-anxiety continuum and other psychiatric variables is linear throughout the range of the continuum. Such a relationship would be consistent with increasing discussions regarding whether or not it could be advantageous (or cost-effective) for clinical services to promote well-being. For example, some professional bodies have called for mental health professionals to jointly focus on increasing well-being and reducing distress (e.g., The British Psychological Society, 2010). Positive Clinical Psychology has endorsed a balanced and equally weighted focus on the positive and negative aspect of life (Wood & Tarrier, 2010, as clarified in Johnson & Wood, in press; Wood & Johnson, 2016). Evidence of a constant linear relationship between the calmness-anxiety continuum and other psychiatric variables would indicate both psychometric and phenomenological continuity and would therefore provide much-needed evidence to substantiate these theoretical arguments. Such evidence would point to the joint importance and usefulness of treating anxiety problems and promoting calmness since calmness and anxiety reside on the same continuum and changing one therefore changes the other. Psychiatric and psychological interventions that are grounded in this conceptualization would logically be stopped when an individual reports high calmness and could be initiated when an individual begins moving away from high calmness.

3.3. Method

Ethical approval to use two existing datasets was granted to the first author by the University of Hertfordshire, United Kingdom. The datasets used also received ethical approval when they were conducted.

3.3.1. Participants

Two samples were used. Sample 1 comprised 4,138 adolescents aged 13-21 years from Hawai‘i. These individuals took part in the five-year longitudinal Hawaiian High Schools Health Survey (HHSHS) study conducted by the National Center on Indigenous Hawaiian Behavioral Health (NCIHBH). This sample provides a broad spread of ages, ethnicities, socioeconomic status,’ and gender (Andrade et al., 2006; Hishinuma et al., 2001). Participants for the HHSHS study were sampled from five high schools during five consecutive school years (1991/1992 to 1995/1996). The schools were selected from both urban and rural areas to obtain a representative sample of adolescents residing in Hawaii. Students who provided assent completed the survey in their classrooms under the supervision of their teachers. Parents of students younger than 18 years old were notified

of the study by mail and given an opportunity to refuse participation. Data collected during the 1992/1993 (*N* = 4,164), 1993/1994 (*N* = 4,182), and 1994/1995 (*N* = 1,433) school years were used in this study. These three school years were used because of the large sample size and inclusion of the variables of interest. This sample completed the STAI and measures of depression, aggression, and substance misuse. There was some missing demographic information and incomplete questionnaire responses (see Andrade et al., 2006; Hishinuma et al., 2001). We multiply imputed missing data as best practice (discussed below). However, the dataset authors report that there were no significant differences on the STAI between participants who had intact versus missing data (Hishinuma et al., 2001).

Sample 2 comprised 1,824 British pregnant women aged 16-43 years (*M* = 26.8 years) from the Cambridge Prenatal Screening Study (CPSS). The purpose of the study was to examine the knowledge, attitudes, anxieties, and experiences of pregnant women booked for antenatal care at hospitals with differing screening policies. Participants were recruited from nine District hospitals within 60 miles of Cambridge (United Kingdom; UK) between 1990-1991. They provided information by telephone and mail interviews. The response rate was 53%. This sample only completed the STAI and was included in order to replicate our structural analyses. Trait STAI items were completed at 12 weeks pregnant; State STAI items were completed at 35 weeks pregnant and 6 weeks post-natal. Full demographic details and procedures are reported elsewhere (Green, Statham & Snowdon, 1996).

3.3.2. Measures

State-Trait Anxiety Inventory (STAI; Spielberger et al., 1970). The STAI is described as a measure of anxiety problems. It consists of 40 items. Trait anxiety is seen as a relatively stable individual difference in the tendency to respond to situations perceived as threatening with elevation in state anxiety (Spielberger et al., 1970). State anxiety is conceptualised within the STAI as a transitory emotional state characterized by subjective, consciously perceived feelings of tension and apprehension and heightened autonomic arousal. The trait scale comprises thirteen negatively worded items (e.g., "I worry too much over something that really doesn't matter") and seven positively worded items (e.g., "I am cool, calm and collected"). Trait items are rated on a 4-point frequency scale based on "how you generally feel." The state scale comprises ten negatively worded items (e.g., "I feel anxious") and 10 positively worded items (e.g., "I feel calm"). State items are rated on a 4-point frequency scale, with instructions asking readers to rate based on "how you feel right now, that is, at this moment." State and trait scales demonstrate excellent internal

consistency (average Cronbach's αs > .89) and the trait scale has evidenced excellent test–retest reliability at multiple time intervals (average *r* = .88; Barnes, Harp & Jung, 2002). The current samples demonstrated similar internal consistency values (Cronbach's α = .89-.90). The state scale demonstrates lower temporal stability (average *r* = .70), as would be expected given the nature of the construct (Barnes et al., 2002). Trait and state scales have evidenced adequate convergent validity with other measures of state and trait anxiety (Spielberger, 1989) and discriminant validity from, for example, aggression (*r* = .38) and substance use (*r* = .19; Knight, Waal-Manning & Spears, 1983). The STAI has been validated with a range of ethnic groups (e.g., Boeke, Duivenoorden & Bonke, 1984; Canals, Marti-Henneberg, Fernandez-Ballart, Clivillle & Domenech, 1992; Vautier & Pohl, 2009). The STAI demonstrated excellent internal consistency (α = 89).

Across both samples and for the State and Trait scales, items representing anxiety were generally positively skewed and items representing calmness were either negatively skewed or showed a normal distribution.

Distribution of distributions of average responses to the STAI State items

		Number of days per week symptoms experienced, on average			
Sample	Items	Not at all	Somewhat	Moderately so	Very much so
Hawaiian	Calmness	1199	1333	1213	521
	Anxiety	2204	984	431	246
British	Calmness	292	535	393	139
	Anxiety	691	364	201	97

Distribution of distributions of average responses to the STAI Trait items

		Number of days per week symptoms experienced, on average			
Sample	Items	Almost never	Sometimes	Often	Almost always
Hawaiian	Calmness	826	1366	1270	315
	Anxiety	1223	1558	677	331
British	Calmness	574	537	554	105
	Anxiety	648	809	222	91

Centre for Epidemiological Studies-Depression (CES-D; Radloff, 1977). The CES-D is one of the most frequently used self-report measures of depressive experiences (Santor, Gregus & Welch, 2006). Responses capture the frequency of feelings and behaviours over the past seven days and are rated on a 4-point scale ranging from 0 (rarely or none of the time) to 3 (most or all of the time). The CES-D contains 20 items that are summed so that scores have a potential range from 0 to 60, with higher scores indicating greater frequency of depressive experiences (Radloff, 1977). The CES-D has been shown

to have good psychometric properties, including high internal consistency in community and psychiatric populations (Cronbach's αs = .85 - .90; Ensel, 1986; Radloff, 1977; Roberts, 1980); convergent validity with other popular measures of depressive experiences such as the Patient Health Questionnaire-9 (r = .85; Amtmann et al., 2014) and Beck Depression Inventory-II (r = .86; Shean & Baldwin, 2008); and discriminant validity from, for example, aggression (r = .44) and substance use (r = .24; Makini et al., 1996). The CES-D demonstrated excellent internal consistency in the current sample (Cronbach's α = .88). A cutoff score of 16 has been found to have sensitivity and specificity rates of 86.7 and 76.6 for identifying depressed individuals, whereas a cutoff score of 21 has a sensitivity and specificity rate of 73.0 and 96.1 (Shean & Baldwin, 2008). The CES-D has been validated with a range of ethnic groups (e.g., Andrade et al., 2006; Garrison, Addy, Jackson, McKeown, et al., 1991). It demonstrated excellent internal consistency (baseline/1992/1993 school year: α = .89; Time 1/1993/1994 school year: α = .89; Time 2/1994/1995 school year: α = .90).

Substance Abuse Subtle Screening Inventory—Adolescent version (SASSI-A; Miller, 1990). Six items were administered from the Substance Abuse Subtle Screening Inventory-Adolescent (SASSI-A; Miller, 1990) as a brief screen for substance use and impairment and dependency arising from substance use. The SASSI-A has been shown to have good psychometric properties, including acceptable internal consistency in the current sample (Cronbach's α = .74); and discriminant validity from anxiety (r = .19), depression (r = .24), and aggression (r = .33; Makini et al., 1996). The SASSI-A also been shown to concord with a diagnosis of substance abuse and dependency on the Diagnostic Interview Schedule for Children (Nishimura et al., 2001), predict counselor DSM-III diagnoses for dually diagnosed adolescent inpatients (Piazza, 1996), and predict adolescent chemical dependency (Risberg, Stevens, & Graybill, 1995). The SASSI-A has been validated with a range of ethnic groups (Nishimura et al., 2001). It demonstrated good internal consistency (baseline/1992/1993 school year: α = .88; Time 1/1993/1994 school year: α = .88; Time 2/1994/1995 school year: α = .86).

Braver Aggressiveness Dimension Scale (BADS; Braver, Fogas, Sandler & Volchik, 1986). The BADS is a 14-item abbreviated self-report measure of child and adolescent aggression. It was derived from the longer Youth Self-Report scales (YSR: Achenbach, 1991), the self-report version of the Child Behaviour Checklist. Items selected for the BADS were those items from the YSR which were significantly more likely to be endorsed by clinically-diagnosed, conduct disordered children and adolescents. The BADS has good psychometric properties, including good internal consistency in the current

sample (Cronbach's α = .85); one year test-retest stability (*r* = .61); and discriminant validity from anxiety (*r* = .38), depression (*r* = .44), and substance use (*r* = .33), with which it shares only moderate correlation (Makini et al., 1996). The BADS demonstrated good internal consistency (baseline/1992/1993 school year: α = .85; Time 1/1993/1994 school year: α = .70; Time 2/1994/1995 school year: α = .64).

Missing Data

There were substantial amounts of missing data in the Hawaiian dataset. 6.4% of all values were missing for the 1992/1993 school year, 52.2% of all values were missing for the 1993/1994 school year, and 22.53% of all values were missing for the 1994/1995 school year. The missingness was not completely at random (MCAR). We addressed this potential problem by multiply imputing missing data on all variables at the item level using SPSS version 21.0 (IBM Corp, 2012). Multiple imputation (MI) is increasingly advocated as the optimal approach for dealing with missing data (Graham, 2009; Schafer & Graham, 2002; Shrive, Stuart, Quan & Ghali, 2006). When MI has been compared with alternative methods of handling incomplete data (e.g., single imputation methods, complete-case analyses, maximum likelihood approaches), it has been shown to generate less biased estimates that have more statistical efficiency (e.g., Crawford, Tennstedt & McKinlay, 1995; Donders, van der Heijden, Stijnen & Moons, 2006; Liu & Gould, 2002; Tang, Song, Belin & Unutzer, 2005). There is also evidence indicating that MI performs well across different circumstances, such as small samples, very large multiple regressions, and when there are large amounts of missing data (Graham & Schafer, 1999).

MI works by generating plausible missing values multiple times based on the distribution of the observed data. Random components are incorporated into these estimated values to reflect their uncertainty. This procedure creates a set of ‘‘complete’’ data sets with no missing values. Analyses are then run separately on each data set, and the results are pooled across datasets using multiple imputation combining rules (Enders, 2010; Graham, 2009). The purpose of MI is not to obtain the individual values themselves but to estimate unbiased parameter estimates of the data set as a whole (Graham, 2009). We followed recommendations to match the number of imputations to the fraction of missing information because progressively larger numbers of imputed datasets are needed to maximize power in subsequent significance testing (Bodner, 2008; Graham, Olchowski & Gilreath, 2007; White et al., 2011).

Statistical Analysis

Analyses were conducted using SPSS version 21.0 (IBM Corp, 2012) and R (R Development Core Team, 2009). Confirmatory Factor Analysis (CFA) was performed

using the R lavaan package, version 0.5-18 (Rosseel, 2012). Positively worded STAI items were reverse scored for all analyses. Three CFA models were tested using full information maximum likelihood (ML) estimation. Model 1 was the standard two factor model involving separate negatively worded items ("anxiety present") and positively worded items ("anxiety absent"), which were allowed to correlate. Model 2 was a single factor model with all items loading on a single factor. Model 3 featured a single substantive anxiety/calmness factor, but all positively worded items were allowed to cross-load onto a second methodological artefact factor which takes into account additional residual inter-correlation between positively worded items. The three CFA models were estimated separately for the state and trait subscales and the two samples to ensure that findings were not specific to a particular form of the STAI or sample.

Acceptable fit was operationalized as Root Mean Squared Error of Approximation (RMSEA) $\leq$ .08, Comparative Fit Index (CFI) $\geq$ .90, Tucker Lewis Index (TLI) $\geq$ .90. Good fit was operationalized as RMSEA $\leq$ .06, CFI $\geq$ .95, and TLI $\geq$ .95 (Hu & Bentler, 1999). Competing models were compared using (i) Akaike's Information Criterion (AIC), which tests the relative fit of competing models after adjusting for parsimony (lower AICs indicate less information loss and thus a superior model), (ii) CFI, using a .002 cutoff (Meade, Johnson & Braddy, 2008), and (iii) Bayesian Information Criterion (BIC), where lower BIC statistics suggest better fit. Although the AIC and BIC share the same goodness-of-fit term, the penalty term of BIC is potentially much more stringent than the penalty term of AIC so BIC tends to choose fitted models that are more parsimonious than those favored by AIC. The AIC, CFI, and BIC model comparison indices have the advantage of being less compromised by large sample sizes when compared to the chi-squared and chi-squared difference statistics (see Cheung & Rensvold, 2002; Meade et al., 2008).

As STAI data are ordinal, we conducted a robustness check of our CFA results by replicating them using mean- and variance-adjusted weighted least squares (WLSMV) estimation. WLSMV estimation has been found to result in unbiased parameter and standard error estimates, and acceptable type-I error rates for structural equation modelling with (skewed) ordinal variables (Flora & Curran, 2004). Competing models were compared using change in model fit according to CFI and RMSEA (Chen et al., 2008).

Hierarchical ordinary least squares (OLS) regressions were used to explore linear and nonlinear relationships between STAI trait scores, treated as a single factor (all items summed to produce a total trait score), and outcome variables. In each analysis, Step 1 involved fitting a model whereby STAI trait scores had a linear relationship with each outcome variable measured at the same time (1992/1993 school year), or measured at

follow-up 1 or 2 years later (1993/1994, 1994/1995 school years), whilst controlling for scores on the outcome variable at baseline (hence it was the change in outcome that we were predicting). Steps 2 and 3 tested whether adding a nonlinear term (squared and cubed STAI trait total scores) made a significant improvement to the amount of variance explained. Improvement in model fit was based on ΔR^2. Statistically significant deviations from linearity were graphed in order to visually display relationships, using unstandardized regression coefficients. This also clarified whether nonlinearity was substantive. The state score, by its very nature, was not expected to reliably predict changes in outcome variables over time and so did not feature in these analyses.

3.4. Results

3.4.1. Comparison of CFA Models

Table 1 shows that across state and trait items and both samples, the two factor model (Model 1) demonstrated an improvement in fit over the single factor model (Model 2). This replicates previous findings. However, Model 3 subsequently outperformed the traditional two factor model, suggesting that when shared method bias amongst positively worded items is controlled for, a single factor underlies the STAI items. The AIC and BIC statistics, which account for model complexity, and the change in model fit according to CFI and RMSEA (Chen et al., 2008), all indicated superiority of Model 3 as hypothesized. Overall fit was acceptable for Model 3 in all instances, with the exception of the state items in the Hawai‘i sample, where fit indices fell slightly below our criteria. Nonetheless the adjusted model demonstrated better fit than the traditional two factor model and was therefore favoured.

Our robustness check employing WLSMV estimation corroborated these results (see Supplementary material for Chapter 3). Again, the two factor model demonstrated an improvement in fit over the single factor model in all instances. Model 3 outperformed the traditional two factor model in the British sample and performed similarly in the Hawai‘i sample. Overall fit was acceptable for Model 3 in the majority of instances. Taken together, the CFA results using ML and WLSMV estimation support the superiority of Model 3 (a calmness-anxiety continuum) over the traditional two factor model (separate “anxiety present” and “anxiety absent” factors), suggesting that when shared method bias amongst positively worded items is controlled for, a single factor underlies STAI state and trait scales.

3.4.2. Exploration of Linear and Nonlinear Relationships with Outcome Variables

A series of regression analyses were conducted in the Hawai‘i sample to explore the form of the relationship between STAI trait scores and outcome variables over time

(Table 2). Three of the nine regression models showed statistically significant nonlinear relationships for Step 2. However, the squared term accounted for very little additional variation above and beyond the linear main effect (2.4%, 1.6%, 0.1%). Two of the nine regression models showed statistically significant nonlinear relationships for Step 3. However, again, the cubed term accounted for very little additional variation (0.2%, 0.2%). Thus, in all cases the nonlinear term failed to make any substantive improvement to the original linear model and these results provide only very weak evidence of a nonlinear relationship.

Potential nonlinearity was further explored by graphing statistically significant nonlinear relationships (Figure 1). The graphs reveal only subtle variation away from perfect linearity. Given the large proportion of missing data, we conducted a robustness check of our regression analyses using complete cases (see Supplementary material for Chapter 3). These results were almost identical (in terms of ΔR2 values), again finding that baseline levels on the calmness-anxiety continuum have a near linear relationship with changes in depression, aggression, and substance misuse over time (regression coefficients are larger than those in Table 2 as multiple imputation estimates relationships more conservatively).

Overall, the regression analyses and graphs provide evidence that the calmness-anxiety continuum (STAI trait total scores) has a near linear relationship with outcome variables over time. Any nonlinearity appears to be of statistical but not practical or clinical significance. These results support our prediction that moving from anxiety to calmness on the STAI provides an equal decrease in risk for several other psychological problems, irrespective of position on the calmness-anxiety continuum.

Table 1. Comparison of Three Maximum Likelihood CFA Models in two Independent Samples.

Model	Model fit						
	X^2	*df*	AIC	BIC	TLI	CFI	RMSEA
British sample, trait items (N = 1824)							
1. Two factor	1521.088*	169	105393.044	105618.904	.938	.945	.066
2. Single factor	2297.308*	170	106167.264	106387.615	.903	.913	.083
3. Single factor, method variance factor	1399.848*	163	105283.803	105542.716	.941	.950	.064
British sample, state items (N = 1824)[1]							
1. Two factor	1778.852*	169	77214.564	77429.490	.928	.936	.083
2. Single factor	2773.269*	170	78206.982	78416.665	.884	.896	.105
3. Single factor, method variance factor	1464.346*	160	76918.058	77180.162	.938	.948	.076
Hawai‘i sample, trait items (N = 4138)[2]							
1. Two factor	1882.202*	169	150252.538	150503.738	.926	.935	.055
2. Single factor	8382.153*	170	156750.489	156995.562	.649	.686	.119
3. Single factor, method variance factor	1773.737*	163	150156.073	150444.033	.923	.938	.054
Hawai‘i sample, state items (N = 4138)[2]							
1. Two factor	4987.366*	169	160309.299	160562.289	.842	.859	.090
2. Single factor	11325.305*	170	166645.238	166892.060	.636	.674	.136
3. Single factor, method variance factor	4744.901*	160	160084.834	160393.357	.841	.866	.090

[1]STAI at 6 weeks post natal; [2]STAI completed during 1992/1993 school year; analyses are reported to three decimal places for clarity; * $p < .001$

Table 2. Results of Regression Analyses Comparing Linear and Nonlinear Effects of STAI upon Change in Outcome.

Step	Variables	*B*	*SE B*	*β*	ΔR^2
	Depression as outcome				
1992/1993 school year (*N* = 4,069)					
1	Constant	-1.569	.277		
	Total STAI Trait score	.788	.011	.744***	.554***
2	Constant	4.235	.470		
	Total STAI Trait score	.209	.040	.198***	
	Total STAI Trait score squared	.012	.001	.567***	.024***
3	Constant	6.740	.719		
	Total STAI Trait score	-.215	.101	.095*	
	Total STAI Trait score squared	.031	.004	.201***	
	Total STAI Trait score cubed	.000	.000	.116***	.002***
1993/1994 school year (*N* = 4,101)					
1	Constant	12.715	.398		
	1992/1993 CES-D total score	.199	.022	.246***	
	Total STAI Trait score	.159	.024	.164***	.149***
2	Constant	13.141	.668		
	1992/1993 CES-D total score	.222	.023	.242***	
	Total STAI Trait score	.119	.056	.122*	
	Total STAI Trait score squared	.001	.001	.046	.000
3	Constant	12.702	1.011		
	1992/1993 CES-D total score	.223	.023	.025***	
	Total STAI Trait score	.192	.139	.143	
	Total STAI Trait score squared	-.002	.006	.304	
	Total STAI Trait score cubed	.000	.000	.176	.000
1994/1995 school year (*N* = 4,101)					
1	Constant	19.015	.568		
	1992/1993 CES-D total score	.153	.027	.197***	
	Total STAI Trait score	.083	.027	.100***	.081***
2	Constant	18.687	.774		
	1992/1993 CES-D total score	.156	.027	.201***	

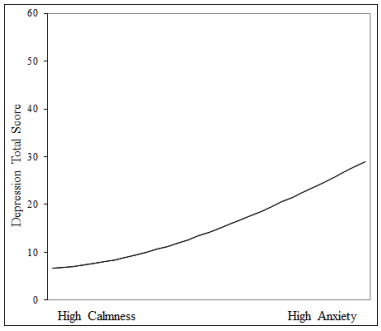
	Total STAI Trait score	.113	.058	.137*	
	Total STAI Trait score squared	-.001	.001	-.042	.000
3	Constant	18.354	1.069		
	1992/1993 CES-D total score	.157	.027	.202***	
	Total STAI Trait score	.169	.135	.205	
	Total STAI Trait score squared	-.003	.006	-.194	
	Total STAI Trait score cubed	.000	.000	.089	.000
	Aggression as outcome				
1992/1993 school year (N = 4,069)					
1	Constant	-.053	.169		
	Total STAI Trait score	.260	.007	.515***	.266***
2	Constant	2.191	.292		
	Total STAI Trait score	.036	.025	.072	
	Total STAI Trait score squared	.005	.000	.461***	.016***
3	Constant	2.747	.447		
	Total STAI Trait score	-.058	.063	-.115	
	Total STAI Trait score squared	.009	.003	.881***	
	Total STAI Trait score cubed	.001	.000	-.247	.001
1993/1994 school year (N = 4,101)					
1	Constant	4.903	.192		
	1992/1993 BADS total score	.341	.017	.367***	
	Total STAI Trait score	.042	.009	.089***	.175***
2	Constant	4.880	.321		
	1992/1993 BADS total score	.341	.018	.367***	
	Total STAI Trait score	.044	.027	.094	
	Total STAI Trait score squared	.000	.001	-.005	.000
3	Constant	4.951	.484		
	1992/1993 BADS total score	.341	.018	.367***	
	Total STAI Trait score	.032	.067	.068	
	Total STAI Trait score squared	.000	.003	.053	
	Total STAI Trait score cubed	.000	.000	-.034	.000
1994/1995 school year (N = 4,101)					
1	Constant	7.743	.256		

	1992/1993 BADS total score	.244	.023	.302***	
	Total STAI Trait score	.021	.010	.051*	.311*
2	Constant	7.416	.342		
	1992/1993 BADS total score	.247	.023	.306***	
	Total STAI Trait score	.053	.025	.129*	
	Total STAI Trait score squared	-.001	.001	-.083	.000
3	Constant	7.452	.484		
	1992/1993 BADS total score	.247	.023	.306***	
	Total STAI Trait score	.047	.063	.114	
	Total STAI Trait score squared	.000	.003	-.050	
	Total STAI Trait score cubed	.000	.000	-.019	.000
	Substance misuse as outcome				
1992/1993 school year (N = 4,069)					
1	Constant	.005	.053		
	Total STAI Trait score	.044	.002	.308***	.095***
2	Constant	.170	.092		
	Total STAI Trait score	.027	.008	.192***	
	Total STAI Trait score squared	.000	.000	.120*	.001*
3	Constant	.469	.141		
	Total STAI Trait score	-.024	.020	-.166	
	Total STAI Trait score squared	.003	.001	.928***	
	Total STAI Trait score cubed	.000	.000	-.474***	.002*
1993/1994 school year (N = 4,101)					
1	Constant	1.111	.053		
	1992/1993 SASSI-A total score	.403	.016	.442***	
	Total STAI Trait score	.005	.002	.036*	.204***
2	Constant	1.125	.085		
	1992/1993 SASSI-A total score	.403	.016	.442***	
	Total STAI Trait score	.003	.007	.025	
	Total STAI Trait score squared	.000	.000	.011	.000
3	Constant	.983	.130		
	1992/1993 SASSI-A total score	.404	.016	.443***	
	Total STAI Trait score	.027	.018	.211	

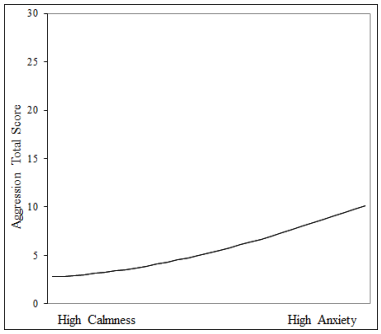
	Total STAI Trait score squared	-.001	.001	-.407	
	Total STAI Trait score cubed	.000	.000	.246	.000
1994/1995 school year (N = 4,101)					
1	Constant	1.714	.067		
	1992/1993 SASSI-A total score	.222	.020	.298***	
	Total STAI Trait score	.005	.002	.051*	.097***
2	Constant	1.650	.094		
	1992/1993 SASSI-A total score	.222	.020	.299***	
	Total STAI Trait score	.012	.007	.112	
	Total STAI Trait score squared	.000	.000	-.064	.000
3	Constant	1.532	.133		
	1992/1993 SASSI-A total score	.223	.020	.300***	
	Total STAI Trait score	.032	.018	.301	
	Total STAI Trait score squared	-.001	.001	-.489	
	Total STAI Trait score cubed	.000	.000	.249	.001

Note: STAI = State Trait Anxiety Inventory; CES-D = Centre for Epidemiological Studies-Depression; BADS = Braver Aggressiveness Dimension Scale; SASSI-A = Substance Abuse Subtle Screening Inventory—Adolescent version; STAI Trait scale completed during 1992/1993 school year; analyses are reported to three decimal places for clarity; *p < .05, ***p < .001.

Depression for
1992/1993 school year



Aggression for
1992/1993 school year



Substance misuse for
1992/1993 school year

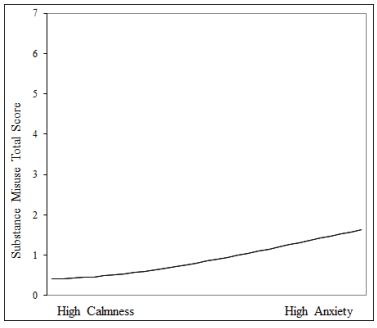


Figure 1. Line graphs plotting unstandardized nonlinear regression lines for statistically significant ΔR^2 values. Total STAI trait scores predict outcome variables at different time points; STAI trait scale completed during 1992/1993 school year.

3.5. Discussion

Our results replicate those of Vautier and Pohl (2009). Like them, we demonstrated that state and trait anxiety, as measured by the STAI, can be understood as continua that range from high calmness to high anxiety. Our analyses were underpinned by Joseph and Wood's (2010) hypothesis that STAI "anxiety absent" items (e.g., "I am cool, calm and collected" [trait anxiety]; "I feel calm" [state anxiety]) assess the presence of calmness, rather than the mere absence of anxiety problems. We provided the first evidence to corroborate this hypothesis. These results have clear implications for the structure and definition of anxiety as they go against the view that anxiety ranges from zero to intense.

We established an anxiety-calmness continuum using the English-language version of the STAI in mixed samples of adults, adolescents, and different ethnic groups from opposite sides of the globe (Hawai'i and the UK) and replicated our results using both ML and WLSMV estimation. It was important to clarify that anxiety-calmness continua existed across diverse circumstances because the mechanisms underlying anxiety may differ across groups (e.g., Field & Lester, 2010; Kirmayer, Young & Hayton, 1995; Manson, 1996), which could lead to misleading artifacts of non-invariant measurement (e.g., item content or wording that is biased against a given group). The use of a large dataset that is representative of adolescents residing in Hawaii minimized the likelihood of systematic sampling bias, which could have been introduced had we used a purely community or clinical sample.

Previous factor-analytic evidence for separate "anxiety present" and "anxiety absent" factors likely arose because there is additional common variance between positively worded items that is unrelated to the underlying latent variable. Based on item content, the STAI state and trait continua can be reconceptualized as calmness-anxiety continua (Joseph & Wood, 2010; Vautier & Pohl, 2009). This conceptualization is more intuitive than the "anxiety present" and "anxiety absent" conceptualization. We note that it is still appropriate to continue to reverse score positively worded STAI items in order to produce total state and trait scores. However, in light of the present findings, STAI users are advised to interpret total state and trait scores as an indication of anxiety problem severity that is based on a combination of the presence/absence of anxiety problems and the presence/absence of calmness, understanding that as anxiety problems increase, calmness decreases (and vice-versa).

As the STAI shows high convergent validity with other leading measures of anxiety problems (Barnes et al., 2002; Knight et al., 1983; Spielberger, 1989), a calmness-anxiety continuum may be apparent in other scales that contain factors that consist of entirely

positively and negatively worded items. Scales which measure anxiety or calmness, but which do not contain a mixture of positively and negatively worded items, are presumably measuring one half of the continuum. Evidence of a calmness-anxiety continuum suggests that existing research into anxiety problems will have relevance for the field of calmness and relaxation research, and vice-versa, and that studying anxiety or calmness separately may be unnecessarily duplicating research effort (Joseph & Wood, 2010).

This is the first study to characterise the form of the relationship between the calmness-anxiety continuum and other psychiatric variables. These analyses make the clinical importance of this topic more apparent and explicit than merely examining whether a continuum exists as Vautier and Pohl (2009) did. Our results demonstrate that baseline levels on the calmness-anxiety continuum have a near linear relationship with outcome variables measured at the same time and one and two years later. Thus, there is no intrinsic way to demarcate problematic degrees of anxiety (or beneficial degrees of calmness) based on how anxiety or calmness are related to other psychiatric variables. That these results were apparent over time suggests that moving along the continuum towards high calmness provides continuous and long-term protection against experiencing other psychological problems.

The present study provides evidence to substantiate calls by professional bodies (e.g., The British Psychological Society, 2010), the Positive Clinical Psychology movement (e.g., Wood & Tarrier, 2010, as clarified in Johnson & Wood, in press; Wood & Johnson, 2016), and many clinicians, for clinical services to adopt a broader focus that jointly involves reducing distress and increasing well-being. Our results point to the usefulness of early intervention and prevention (when people begin to move away from high calmness) and instilling resilience (by providing interventions to move people towards high calmness). Fostering high levels of calmness would mean that individuals have further to go before they reach high levels of anxiety. Psychiatric and psychological interventions that are grounded in a continuum conceptualization would logically be stopped when an individual reports high calmness. Patients, service commissioners and others may of course want interventions to stop at the point of problem absence. However, the present results provide an evidence-base to inform a collaborative discussion around when to stop treatment and the advantages and disadvantages of doing so. This process may already be happening when clinicians construct relapse prevention plans and offer booster sessions with people who have finished treatment.

We are hopeful that our findings could help support a case for publically funded clinical services to accept the promotion of well-being into their remit. That well-being interventions often fall outside the focus of publically funded clinical services means that

efforts to help the public address this need (e.g., through self-help books, self-development courses, and other "interventions") are often offered by unaccredited and untrained individuals. We find this concerning, as we strongly believe that clinical interventions should be targeted, theory-driven, evidence-based, and provided by suitably qualified, ethically-practicing professionals.

Our results could be extended in a number of ways. Because cost considerations have a substantial impact on service delivery and often outweigh theoretical or moral arguments, the current results need to be accompanied by a comprehensive economic cost value analysis which tests the implications of the continuum conceptualization and when it is most cost effective to stop interventions. It is obviously less cost-effective in the short-term to stop interventions at the point of well-being rather than mere problem absence, or to start interventions when people begin to move away from high calmness but before a severe psychological problem becomes manifest. However, this approach may prove to be the most cost-effective solution overall if it provides long-term protection from other problems, especially amongst high risk groups. Research is also needed to characterise the form of the relationship between the calmness-anxiety continuum and other psychiatric variables in relation to all age groups who complete the STAI.

We also note that the current findings were limited in not accounting for random and systematic measurement error (Barrett & Russell, 1998; Green et al., 1993). Future research investigating the calmness-anxiety continuum could account for intraindividual mood variation and measurement error by taking repeated (e.g., daily) continuous measures of mood using different response formats.

3.6. Supplementary Material for Chapter 3

Table S1. Comparison of Three Mean- and Variance-Adjusted Weighted Least Squares CFA Models in two Independent Samples.

Model	Model fit					
	X^2	*df*	AIC	TLI	CFI	RMSEA
British sample, Trait items (N = 1824)[a]						
1. Two factor	1737.857*	169		.948	.954	.071
2. Single factor	4027.596*	170		.872	.886	.112
3. Single factor, method variance factor	1681.756*	163		.948	.955	.071
British sample, State items (N = 1824)[a]						
1. Two factor	3445.255*	169		.905	.916	.118
2. Single factor	6323.236*	170		.823	.842	.161
3. Single factor, method variance factor	2792.752*	160		.920	.932	.109
Hawai‘i sample, Trait items (N = 4138)[b]						
1. Two factor	3448.257*	169		.937	.944	.076
2. Single factor	13518.638*	170		.744	.771	.152
3. Single factor, method variance factor	3370.175*	163		.936	.945	.076
Hawai‘i sample, State items (N = 4138)[b]						
1. Two factor	8547.603*	169		.891	.903	.118
2. Single factor	18207.190*	170		.766	.790	.173
3. Single factor, method variance factor	8480.062*	160		.885	.903	.121

[a]STAI at 6 weeks post natal; [b]STAI completed during 1992/1993 school year; analyses are reported to three decimal places for clarity; * $p < .001$.

Table S2. Results of Regression Analyses Based on Complete Cases Comparing Linear and Quadratic Effects of STAI upon Change in Outcome.

Step	Variables	*B*	*SE B*	*β*	ΔR^2
	Depression as outcome				
1992/1993 school year (*N* = 3,759)					
1	Constant	-1.255	.271		
	Total STAI Trait score	.770	.011	.747***	.558***
2	Constant	4.715	.448		
	Total STAI Trait score	.154	.039	.150***	
	Total STAI Trait score squared	.013	.001	.622***	.030***
3	Constant	7.200	.666		
	Total STAI Trait score	-.285	.096	-.277***	
	Total STAI Trait score squared	.033	.004	1.593***	
	Total STAI Trait score cubed	.000	.000	-.576***	.003***
1993/1994 school year (*N* = 1,834)					
1	Constant	3.351	.447		
	1992/1993 CES-D total score	.348	.028	.366***	
	Total STAI Trait score	.236	.028	.244***	.327***
2	Constant	5.564	.775		
	1992/1993 CES-D total score	.321	.028	.338***	
	Total STAI Trait score	.021	.068	.021	
	Total STAI Trait score squared	.005	.001	.253***	.004***
3	Constant	4.728	1.141		
	1992/1993 CES-D total score	.324	.029	.340***	
	Total STAI Trait score	.171	.165	.176	
	Total STAI Trait score squared	-.002	.007	-.105	
	Total STAI Trait score cubed	.000	.000	.213	.000
1994/1995 school year (*N* = 870)					
1	Constant	5.068	.703		
	1992/1993 CES-D total score	.332	.046	.334***	
	Total STAI Trait score	.185	.046	.184***	.239***
2	Constant	6.622	1.204		
	1992/1993 CES-D total score	.311	.047	.313***	

	Total STAI Trait score	.031	.107	.031***	
	Total STAI Trait score squared	.004	.002	.176	.002
3	Constant	6.140	1.797		
	1992/1993 CES-D total score	.313	.048	.315***	
	Total STAI Trait score	.117	.261	.117	
	Total STAI Trait score squared	.000	.012	-.021	
	Total STAI Trait score cubed	.000	.000	.116	.000
	Aggression as outcome				
1992/1993 school year (N = 3,812)					
1	Constant	.117	.134		
	Total STAI Trait score	.257	.007	.523***	.274***
2	Constant	2.215	.277		
	Total STAI Trait score	.040	.024	.081	
	Total STAI Trait score squared	.005	.000	.460***	.016***
3	Constant	2.826	.411		
	Total STAI Trait score	-.069	.059	-.140	
	Total STAI Trait score squared	.010	.003	.965***	
	Total STAI Trait score cubed	.000	.000	-.299*	.001*
1993/1994 school year (N = 1,889)					
1	Constant	.781	.199		
	1992/1993 BADS total score	.543	.020	.563***	
	Total STAI Trait score	.056	.010	.119***	.402***
2	Constant	1.356	.341		
	1992/1993 BADS total score	.538	.020	.558***	
	Total STAI Trait score	-.004	.030	-.008	
	Total STAI Trait score squared	.001	.001	.135*	.001*
3	Constant	1.371	.500		
	1992/1993 BADS total score	.538	.020	.558***	
	Total STAI Trait score	-.007	.073	-.014	
	Total STAI Trait score squared	.001	.003	.149	
	Total STAI Trait score cubed	.000	.000	-.008	.000
1994/1995 school year (N = 901)					
1	Constant	1.275	.320		

	1992/1993 BADS total score	.518	.033	.501***	
	Total STAI Trait score	.05	.02	.10*	.311*
2	Constant	1.038	.545		
	1992/1993 BADS total score	.520	.034	.503***	
	Total STAI Trait score	.073	.049	.148	
	Total STAI Trait score squared	-.001	.001	-.054	.000
3	Constant	.884	.809		
	1992/1993 BADS total score	.520	.034	.504***	
	Total STAI Trait score	.101	.119	.204	
	Total STAI Trait score squared	-.002	.005	-.184	
	Total STAI Trait score cubed	.000	.000	.077	.000
	Substance misuse as outcome				
1992/1993 school year (N = 3,804)					
1	Constant	.017	.051		
	Total STAI Trait score	.041	.002	.299*	.090***
2	Constant	.218	.087		
	Total STAI Trait score	.020	.008	.146*	
	Total STAI Trait score squared	.000	.000	.159*	.002***
3	Constant	.462	.130		
	Total STAI Trait score	-.023	.019	-.171	
	Total STAI Trait score squared	.002	.001	.881***	
	Total STAI Trait score cubed	.000	.000	-.428*	.002*
1993/1994 school year (N = 1,826)					
1	Constant	.161	.057		
	1992/1993 SASSI-A total score	.619	.021	.583***	
	Total STAI Trait score	.010	.002	.080***	.374***
2	Constant	.218	.097		
	1992/1993 SASSI-A total score	.618	.021	.582	
	Total STAI Trait score	.004	.009	.033	
	Total STAI Trait score squared	.000	.000	.049	.000
3	Constant	.040	.142		
	1992/1993 SASSI-A total score	.619	.021	.583***	
	Total STAI Trait score	.037	.021	.287	

	Total STAI Trait score squared	-.001	.001	-.537	
	Total STAI Trait score cubed	.000	.000	.351	.001
1994/1995 school year (N = 862)					
1	Constant	.307	.089		
	1992/1993 SASSI-A total score	.471	.036	.413***	
	Total STAI Trait score	.015	.004	.124***	.213***
2	Constant	.275	.151		
	1992/1993 SASSI-A total score	.471	.036	.413***	
	Total STAI Trait score	.019	.014	.152	
	Total STAI Trait score squared	-.007	.000	-.029	.000
3	Constant	-.071	.224		
	1992/1993 SASSI-A total score	.472	.036	.414***	
	Total STAI Trait score	.082	.033	.655*	
	Total STAI Trait score squared	-.003	.001	-1.171*	
	Total STAI Trait score cubed	.000	.000	.676*	.004*

Note: STAI = State Trait Anxiety Inventory; CES-D = Centre for Epidemiological Studies-Depression; BADS = Braver Aggressiveness Dimension Scale; SASSI-A = Substance Abuse Subtle Screening Inventory—Adolescent version; STAI Trait scale completed during 1992/1993 school year; analyses are reported to three decimal places for clarity; *p < .050, ***p < .001.

Table S3. Table Summarizing Which State Trait Anxiety Inventory (STAI) Items Measure Anxiety and Calmness.

	Trait items		State items	
	Anxiety	Calmness	Anxiety	Calmness
Item 1		I feel pleasant		I feel calm
Item 2	I tire quickly			I feel secure
Item 3	I feel like crying		I am tense	
Item 4	I wish I could be as happy as others seem to be		I am regretful	
Item 5	I am losing out on things because I can't make up my mind soon enough			I am at ease
Item 6		I feel rested	I feel upset	
Item 7		I am "calm, cool, and collected"	I am presently worrying over possible misfortunes	
Item 8	I feel that difficulties are piling up so that I cannot overcome them			I feel rested
Item 9	I worry too much over something that really doesn't matter		I feel anxious	
Item 10		I am happy		I feel comfortable
Item 11	I am inclined to take things hard			I feel self-confident
Item 12	I lack self-confidence		I feel nervous	
Item 13		I feel secure	I am jittery	
Item 14	I try to avoid a crisis or difficulty		I feel "high strung"	
Item 15	I feel fed up			I am relaxed
Item 16		I am content		I feel content
Item 17	Some unimportant thought runs through my mind and bothers me		I am worried	
Item 18	I take disappointments so keenly that I can't put them out of my mind		I feel over-excited and rattled	

Item 19		I am a steady person	I feel joyful
Item 20	I become tense and upset when I think about my present concerns		I feel pleasant

CHAPTER 4

4. Characterizing Self-Injurious Cognitions Part I: Development of the Suicide Attempt Beliefs Scale (SABS) and the Nonsuicidal Self-Injury Beliefs Scale (NSIBS)

4.1. Abstract

Self-injurious cognitions (SICs) are cognitions about deliberately injuring oneself (self-injurious behavior; SIB). Existing measures of SICs provide varying content coverage, pointing to a lack of consensus; or do not elucidate why people engage in SIB specifically, rather than performing another behavior. Such measures have also not tested whether it is possible or useful to conceptualise suicide attempts (SA) and nonsuicidal self-injury (NSSI), two forms of SIB, as separate constructs. We developed the Suicide Attempt Beliefs Scale (SABS) and Nonsuicidal Self-Injury Beliefs Scale (NSIBS) to address these issues. A companion article (Part II) describes the validation of these instruments. We tested whether SA and NSSI are considered separate constructs by asking participants to complete a large item pool separately in relation to SA and NSSI. Different ratings of exactly the same items for the two behaviors justified the development of separate scales. A series of exploratory and confirmatory factor analyses across six large, heterogeneous samples (total N = 3,313) revealed that the SABS consists of seven correlated factors, the NSIBS consists of ten correlated factors, and these instruments are separate. SA and NSSI are characterised by some very similar beliefs and some beliefs that are specific to SA or NSSI. Both instruments contain factors that describe how SA or NSSI relates to oneself and others. Results indicate that SA and NSSI are similar but distinct phenomena, supporting the use of separate terminology and definitions of SA and NSSI, and indicating that SA and NSSI need to be separated in research designs.

Chapter 4 will be submitted for peer-review to *Psychological Assessment* as: Siddaway, A. P., Wood, A. M., O'Carroll, R. E., & O'Connor, R. C. Characterizing Self-Injurious Cognitions Part I: Development of the Suicide Attempt Beliefs Scale (SABS) and the Nonsuicidal Self-Injury Beliefs Scale (NSIBS).

4.2. Introduction

When a clinician encounters someone who is contemplating or has engaged in some form of self-injurious behavior (SIB), they are keen to assess the precise details of exactly *what* that person is thinking, and, where relevant, in the details of the SIB itself. This information can be coupled with theory and other clinical and contextual information to make individualised predictions about risk and to formulate targeted therapeutic interventions. An individual's cognitions specifically *about* SIB – their self-injurious cognitions (SICs) – convey what SIB *means* to that person. For example, "People think that my suicide attempt(s) are selfish" or "Nonsuicidal self-injury helps me escape negative emotions." Understanding the specific content of an individual's SICs is absolutely vital because people contemplate SIB far more than they take action (e.g., Fergusson, Beautrais & Horwood, 2003; Ten Have et al., 2009; Kessler, Berglund, Borges, Nock & Wang, 2005; Nock, Prinstein & Sterba, 2009), and because clinical services seek to avoid SICs translating into action. Cognition and behaviour are equally important, although the most accurate way to conceptualise SIB is probably in terms of cognitions about SIB that for some people, at some times, manifest behaviorally.

Comprehensive and reliable measurement of SICs is a critical foundation for understanding, researching, preventing, and treating SIB and, after over 50 years of research, one might assume that the measurement and conceptualisation of SICs would be clear and agreed. Unfortunately, however, this does not appear to be the case, as we outline below. This article (Part I) describes the development of two new multidimensional self-report measures of SICs, the Suicide Attempt Beliefs Scale (SABS) and the Nonsuicidal Self-Injury Beliefs Scale (NSIBS), which were designed to extend and improve on existing measures. We developed these scales in an attempt to characterise SICs in much greater detail than has ever been done before, and to test whether the same cognitions characterise suicide attempts (SA) and nonsuicidal self-injury (NSSI). A companion article (Part II; Siddaway, Wood, O'Carroll & O'Connor, in submission) describes the validation of the SABS and NSIBS.

Defining Self-Injurious Behavior

SIB is uniquely characterised by cognitions and behavior that involve deliberately physically injuring oneself. This conceptualisation distinguishes SIB from: (1) incidental physical injuries sustained as part of generally accepted sociocultural practices (e.g., body piercing, tattooing, bikini-line waxing, neck elongation, facial scarification; Clarke & Whittaker 1998; Favazza, 1989); (2) behaviors which immediately but unintentionally result in physical injury (e.g., accidentally touching something hot; drink driving; skin picking; nail

biting; hair-pulling); (3) behaviors in which physical injury may sometimes be a delayed or unanticipated side-effect (e.g., smoking, over-eating, binge drinking, or unprotected sex); and (4) behaviors which are self-defeating/self-destructive but which do not lead to physical injury (e.g., remaining in an emotionally abusive relationship).

Two forms of SIB can be distinguished depending on the presence or absence of any reported or inferred intent to kill oneself (Linehan et al., 2006; Muehlenkamp, 2014; Nock et al., 2008; Silverman et al., 2007a, b). A "suicide attempt" (SA) is when someone intentionally physically injures themselves because they want and expect to kill themselves. "Suicide" may or may not be a consequence of a SA (Brown, Henriques, Sosdjan & Beck, 2004; De Moore & Robertson, 1999). "Nonsuicidal self-injury" (NSSI) is when someone intentionally physically injures themselves, but with no desire or intention of killing themselves or being dead[3]. SA and NSSI often co-occur (Klonsky, May & Glenn, 2013; Nock et al., 2006) and relate to one-another (see Table 1). For example, the NSIBS *Anti-suicide* factor (described later) taps perceptions that engaging in NSSI is a useful and possible means of avoiding acting on suicidal thoughts (e.g., "NSSI is a compromise instead of killing myself"). Whether or not SAs and NSSI are facets of the same construct is an issue of ongoing controversy (Kapur, Cooper, O'Connor & Hawton, 2013; Muehlenkamp, 2014; Posner, Brodsky, Yershova, Buchanan & Mann, 2014; Silverman, 2016).

Limitations of Existing Measures

A large body of research has attempted to understand the thinking involved in SIB. This has resulted in the development of a range of instruments (for reviews, see Batterham et al., 2014; Brown, 2001; Kodaka, Postuvan, Inagaki & Yamada, 2010; Rothberg & Geer-Williams, 1992; Winters, Myers & Proud, 2002). Broadly speaking, these instruments measure one or more of the following: (1) The presence or absence of thoughts about suicide and/or NSSI (e.g., Linehan et al., 2006; Nock et al., 2007); (2) the characteristics of suicide and/or NSSI thoughts (e.g., frequency, duration, intensity, onset; Beck & Steer, 1991; Claes & Vandereycken, 2007; Gratz, 2001; Klonsky & Glenn, 2009; Linehan & Comtois, 1996; Linehan et al., 2006; Nock et al., 2007; Osman et al., 2001; Sansone et al., 1998; Whitlock et al., 2014); (3) perceptions that are thought to underlie suicidal thoughts and/or SA specifically, such as hopelessness, defeat/entrapment, burdensomeness, unlovability, unbearability, unsolvability, thwarted belongingness, and acquired capability (Beck & Steer,

[3] These issues get at the semantics and idiosyncratic meanings that different individuals associate with different terms.

1993; Gilbert & Allan, 1998; Rudd et al., in preparation; Van Orden et al., 2012); and (4) the reasons, functions, and motivations for engaging in or refraining from SA and/or NSSI (Beck & Steer, 1991; Holden, Kerr, Mendonca & Velamoor, 1998; Kleindienst et al., 2008; Klonsky & Glenn, 2009; Lewis & Santor, 2008; Linehan, Goodstein, Nielsen & Chiles, 1983; Lloyd et al.,1997; Ma & Klonsky, 2013; Matson et al., 1999; Nock et al., 2007; Osman et al., 1998; Osuch et al., 1999; Santa Mina et al., 2006; Turner, Chapman & Gratz, 2014; Whitlock et al., 2014). As we outline next, existing measures of the content of SICs (points 3 and 4 in the list above) are limited in several important ways, each of which points to the need to develop alternative scales.

Specificity

One major shortcoming with many cognitive constructs thought to underlie SAs, such as impulsivity, depression, hopelessness, and perceptions of defeat/entrapment, is that these constructs do not elucidate why people engage in a SA and/or NSSI *specifically* – rather than performing some other behavior. In fact, accumulating evidence demonstrates that these constructs are relevant to a large number of people and psychological problems (e.g., Anestis, Soberay, Gutierrez, Hernandez & Joiner, 2014; Klonsky & May, 2014; Siddaway, Taylor, Wood & Schulz, 2015); the vast majority of individuals reporting these perceptions do not attempt suicide, die by suicide, or engage in NSSI (Goldstein, Black, Nasrallah & Winokur, 1991; Selby et al., 2014). For example, depression is common in people who make a SA but the vast majority of depressed individuals will not attempt to take their own life or die by suicide (Bostwick & Pankratz, 2000).

Focussing on cognitions *about* SIB (SICs) is an approach that may be particularly informative in clarifying why some people, at some times, intentionally physically injure themselves, rather than choosing to use an alternative self-regulatory strategy (e.g., listen to music; go for a walk; problem-solve; ruminate; get drunk) in response to a particular internal or external trigger (e.g., emotional distress; social exclusion). Strong endorsement of SICs may be a specific, defining, and time-varying feature of SIB.

Construct Validity

Measures of the reasons, functions, and motivations[4] for SAs and NSSI are presumably more specific to SIB but are themselves subject to several content-based limitations. Most importantly, there are fifteen or so of these measures and each scale provides different content coverage. This highlights a lack of consensus regarding exactly

[4] These terms and concepts seem to be used fairly interchangeably in the literature.

which cognitions characterise SIB. Furthermore, several of these scales contain multiple items that assess certain types of content but only single items that assess other content domains. This is a potential problem because meaningful content-based factors cannot be identified when only a single relevant marker is included in the item pool (Clark & Watson, 1995; Loevinger, 1957).

Although existing scales cover a range of content, no study to date has included a wide range of SICs in the same item pool and conducted structural analyses to explore exactly which SICs best characterize SIB. Taken together, these issues mean that it is unclear which cognitions reliably characterise SIB and what factor structure best describes SICs. These issues also make it unclear whether any between-study differences observed to date are due to substantive factors, or are simply due to the use of different scales that provide varying and potentially incomplete content coverage.

Prediction, Vulnerability and Maintenance

The ability to predict, prevent, or stop SIB presumably stems from an understanding of SIB vulnerability and maintenance. Unfortunately, however, the phrasing of measures of the reasons for SAs and NSSI may have precluded these measures from being used to investigate how SICs might contribute to SIB vulnerability and maintenance. For example, the Inventory of Statements About Self-Injury (Klonsky & Glenn, 2008) assesses 13 functions of NSSI; each item begins with "When I self-harm, I am …" Likewise, the Inventory of Motivations for Suicide Attempts (Ma & Klonsky, 2013) asks respondents to explain why they attempted suicide using items which begin with "I attempted suicide because I…"

The phrasing used in these measures assumes that people have already attempted suicide or engaged in NSSI, meaning that individuals who are at high risk of SIB but have yet to take action (e.g., as a result of protective internal or external moderator and mediator variables) would not be administered these measures. Asking about reasons for SA or NSSI potentially limits the predictive ability of these measures because the same reasons may be endorsed whether SIB occurred yesterday or 10 years ago, *even though an individual may no longer endorse SICs at all*[5]. The phrasing of these measures also precludes them being used to examine the transition from the first episode of SIB, to repeating SIB, or to examine

[5] The ability to sensitively measure individual differences in SICs and SIB is especially critical because of low base rates; measuring inter-participant variability increases power by increasing the amount of information available for analyses (McClelland & Judd, 1993).

cognitive maintenance factors (e.g., in current versus recovered SIB groups; as a result of psychological therapy modifying cognitions about SIB).

Conceptual Debate

One final concern that can be raised against existing measures of SICs is that they have not directly contributed to the debate regarding whether it is possible or useful to conceptualise SA and NSSI as one or two constructs (see Kapur et al., 2013; Muehlenkamp, 2014; Posner et al., 2014; Silverman, 2016). Clarifying whether and how SICs differ across SA and NSSI has potentially major implications for how we understand, describe, and individualise interventions for SA and NSSI.

The Current Research

The various issues discussed suggest that existing scales may not comprehensively or reliably measure the cognitive content that characterises SA and NSSI. Our goals in this article were to develop comprehensive, theory-driven, multidimensional measures of SICs and to clarify whether the same set of items and/or factors characterise SA and NSSI. We note several key features of our approach.

First, the SABS and NSIBs were designed to measures beliefs about SA and NSSI (rather than thoughts, assumptions, attitudes, reasons, or some other type or level of cognition): items were phrased to reflect beliefs (absolute statements of the way things are regarding SA and NSSI). From a cognitive therapy perspective (e.g., Beck, Rush, Shaw & Emery, 1979), beliefs are conceptualised as relatively enduring personal meanings that confer vulnerability across situations, but which also potentially fluctuate or change due to a range of factors. We developed measures of beliefs about SA and NSSI because we were interested in developing scales that could provide explanatory power beyond current measures (beliefs are thought to explain other types of mental activity; e.g., Beck et al., 1979). Beliefs about SIB potentially transcend individual differences in the phenomenology of SIB (e.g., people can have different automatic thoughts, mental images, or expectations regarding SIB during different episodes of SIB but the beliefs that underline all SIB episodes might be relatively stable).

Second, whilst we were interested in measuring beliefs, we were also keen to develop instruments that could measure differences between individuals and track changes within individuals. The SABS and NSIBS therefore ask how much the reader *currently* agrees with (endorses) each SIC. Third, to enable all potential factors to emerge reliably in structural analyses, we generated a large and varied item pool and included multiple markers to define each potential cognitive content domain. We then conducted factor analyses on the entire

item pool and deleted cross-loading items in order to maximise the discriminant validity of individual subscales.

Fourth, our desire to clarify whether the cognitions that characterise SA are the same as those that characterise NSSI led us to invent a methodology that would directly test this question (more on this below). The outcome of these analyses would justify the development of separate scales or a single scale.

Fifth, when developing these scales, we were acutely aware of the need to balance comprehensiveness against utility. That is, whilst we were keen to develop instruments that measure a wide range of important content, the finished scales would need to be short enough to be appropriate for a population who are often highly distressed and experiencing comorbid psychological problems and cognitive changes such as impaired problem-solving ability (see O'Connor & Nock, 2014, for a review). We achieved this balance by retaining just a subset of items from the final factor structure we obtained in Sample 5, ensuring that each subscale had at least a good level of internal consistency ($\alpha \geq .8$; Clark & Watson, 1995; Nunnally 1978; see Part II).

Last, because substantive validity begins with the creation of a broad item pool that seeks to measure a particular theory (Clark & Watson, 1995; Loevinger, 1957), candidate items were generated to operationalize the hypothesis that three domains of SICs characterise SA and NSSI. Specifically, we predict that the content of SA and NSSI cognitions involves "positive SICs", "negative SICs" and "facilitating SICs" about SIB. A parallel review article elaborates this hypothesis and explores whether the existing literature can be reconceptualised in terms of the three proposed domains of SICs.

"Positive SICs" are cognitions about the idiosyncratic individual and interpersonal perceived advantages of SA and NSSI (e.g., "Attempting suicide changes the way that I am thinking" [individual]; "NSSI helps me fit in with other people" [interpersonal])[6]. The presence and activation of these cognitions are predicted to motivate people to engage in SA and NSSI.

We recognise that there are advantages and disadvantages to doing or not doing any behavior and SIB is no different. Thus, "negative SICs" are cognitions about the idiosyncratic individual and interpersonal perceived disadvantages of SIB (e.g., "NSSI makes my problems

[6] The terminology "positive" is used to convey an individual's *perception*. "Positive" and "negative" terminology are commonly used to describe the valence of psychological phenomena, particularly emotions. "Positive" is not used to imply that SA or NSSI may be perceived as a "helpful" or "constructive" option; only that these behaviours are sometimes seen as personally advantageous and therefore desirable option.

worse" [individual]; "People think that my suicide attempt(s) are selfish" [interpersonal]). The presence and activation of these cognitions are predicted to motivate people to avoid SA and NSSI.

Activation of positive and negative SICs is hypothesised to involve competing motivations to approach and avoid SIB, which is distressing. For example, "I want to cut myself but I know it's bad for me." A third domain of cognitions, "facilitating SICs," are predicted to arise in response to strongly conflicting positive and negative SICs and enable a person to proceed with a behavior that they know to be unconstructive. Facilitating SICs (e.g., "Just do it," "It's OK to cut if I'm really upset") resolve cognitive dissonance by justifying and giving "permission" to self-injure by strengthening positive and/or inhibiting negative SICs.

The three domains of SICs are suggested to be potentially apparent before, during, and after SIB and to interact to maintain SIB. We predict that SA and NSSI are each characterised by the three domains of SICs; however the exact content of SA and NSSI cognitions will potentially differ due to differences in the nature of these behaviours (e.g., frequency, lethality), their consequences, and how each is socially-constructed. Candidate items for the SABS and NSIBS were generated to measure all three domains of SICs.

4.3. Overview of Measure Construction

We highlight several key aspects of scale construction; further details are provided in the Supplementary material for Chapter 4.

Recruitment

The SABS and the NSIBS were developed over six large, heterogeneous samples (total N = 3,313) of people with lived experience of SIB (see Table 1). All samples were recruited online from a broad range of SIB and mental health forums, support websites, and mental health charities worldwide. We specifically recruited individuals with lived experience of any type of SICs or SIB. This sampling strategy enabled us to obtain large, heterogeneous samples and we hoped that this strategy would allow us to develop measures that would be broadly applicable (sampling individuals who vary widely in terms of SICs and SIB e.g., SIB methods, medical lethality and harm, and psychological comorbidity). Sampling individuals outside of clinical settings was also consistent with our theoretical assumption that SICs are continuous traits, rather than discrete, dichotomous entities.

We hoped that anonymous online recruitment from a broad range of sources would provide the best possible opportunity to obtain an accurate and representative understanding of the phenomenology of SICs because: (1) It is well-established that SIB is a highly

stigmatised, often secretive behavior (NICE, 2011), and that the majority of SIB episodes (~70%) do not result in presentation to clinical services (e.g., Hawton Rodham, Evans & Harriss, 2009); (2) some evidence indicates that online recruitment increases SIB reporting by 2-3 times relative to non-anonymous techniques (e.g., Nock et al., 2008); and (3) online participation allowed us to randomise the ordering of our item pools, thereby eliminating an important potential source of error when presenting large numbers of items.

Structural Analyses

The same approach to structural analyses was used in all samples. Parallel analysis was used to determine the number of factors to extract (Velicer, Eaton & Fava, 2000; Zwick & Velicer, 1986) and the results informed maximum-likelihood exploratory factor analyses (EFA) with promax rotation. Various sensitivity analyses were performed to check the robustness of the obtained factor solutions and these consistently yielded similar results.

We were very tentative in our labelling of factors and our deletion of items in Samples 1 and 2 and erred more on the side of sensitivity (administering overly large item pools to potentially encourage meaningful, distinct factors to emerge). By the time Sample 3 was collected, we had used EFA to yield a reliable set of factors for the final scales. Thus, Samples 3-5 focused more on specificity; that is, locating the strongest and most reliable markers for the identified factors.

In each round of factor analysis, the following criteria were adopted to ensure that items would be maximally informative and in order to extract the greatest number of factors that would be well defined and reasonably distinct from one-another. When two items correlated strongly ($\geq$.75), the item with lowest item total correlation was considered to contain redundant information and was deleted (Clark & Watson, 1995). Items with loadings <.40 or which demonstrated reasonably strong loadings on more than one factor (>.3) were eliminated (Clark & Watson, 1995).

Parsimony

We balanced the competing goals of developing scales that are reasonably short (and therefore relatively swiftly completed and clinically useful) but nevertheless reliable and informative (cf. the "attenuation paradox;" Clark & Watson, 1995; Loevinger, 1954) through the following approaches. First, we retained adequate numbers of items in each subscale to achieve at least a good level of internal consistency ($\alpha \geq .8$; Clark & Watson, 1995; Nunnally 1978). Second, as we were keen to develop scales that tap SICs as broadly as possible, rather than simply retaining the highest loading items from each subscale (which would have improved reliability but reduced validity), when subscales contained enough items to allow it,

items were progressively deleted from strongly correlated pairs of items (deleting the item with lowest item total correlation) (Clark & Watson, 1995).

Table 1 Self-Injurious Cognitions and Behavior and Demographic Characteristics by Sample

		Thought capture	Sample 1	Sample 2	Sample 3	Sample 4	Sample 5	Sample 6
N		35	698	436	484	380	664	650
Suicidal thought(s)								
Lifetime presence		33 (94%)	698 (100%)	419 (96%)	469 (97%)	363 (96%)	646 (97%)	628 (97%)
Lifetime frequency								
	1	0 (0%)	1 (0%)	5 (1%)	3 (1%)	11 (%)	6 (1%)	6 (1%)
	2-5	8 (24%)	135 (19 %)	48 (11%)	47 (10%)	55 (%)	78 (12%)	60 (9%)
	5-10	4 (12%)	23 (3%)	56 (13%)	37 (8%)	40 (%)	73 (11%)	70 (11%)
	10-15	1 (3%)	15 (2%)	25 (6%)	30 (6%)	18 (%)	32 (5%)	50 (8%)
	15-20	0 (0%)	19 (3%)	14 (3%)	32 (7%)	21 (%)	43 (7%)	38 (6%)
	20-30	0 (0%)	32 (5%)	28 (6%)	25 (5%)	17 (%)	37 (6%)	26 (4%)
	30+	20 (60%)	473 (68%)	243 (56%)	295 (61%)	201 (53%)	377 (57%)	378 (58%)
Recency								
	Past month	22 (67%)	396 (57%)	255 (59%)	279 (58%)	218 (57 %)	387 (58%)	391 (60%)
	Past year	10 (30%)	168 (24%)	92 (21%)	118 (24%)	73 (19%)	144 (22%)	135 (21%)
	1-2 years ago	1 (3%)	62 (9%)	37 (9%)	28 (6%)	30 (8%)	51 (8%)	39 (6%)
	2+ years ago	-	72 (10%)	35 (8%)	44 (9%)	42 (11%)	64 (10%)	63 (10%)
Suicide attempt(s)								
Lifetime presence		16 (48%)	505 (72%)	268 (62%)	323 (67%)	247 (65%)	427 (64%)	410 (63%)
Lifetime frequency								
	1	3 (9%)	112 (16%)	60 (14%)	55 (11%)	33 (9%)	84 (13%)	81 (13%)
	2-5	9 (27%)	157 (23%)	134 (31%)	167 (34%)	127 (33%)	210 (32%)	197 (30%)
	5-10	1 (3%)	190 (27%)	47 (11%)	52 (11%)	38 (10%)	79 (12%)	71 (11%)
	10-15	0 (%)	12 (2%)	14 (3%)	20 (4%)	23 (6%)	23 (4%)	29 (5%)
	15-20	0 (%)	9 (1%)	6 (1%)	8 (2%)	3 (1%)	9 (1%)	13 (2%)
	20-30	2 (6%)	9 (1%)	5 (1%)	10 (2%)	12 (3%)	7 (1%)	8 (1%)
	30+	1 (3%)	16 (2%)	2 (0%)	11 (2%)	11 (3%)	15 (2%)	11 (2%)
Recency								
	Past month	2 (6%)	54 (8%)	36 (8%)	32 (7%)	28 (7%)	65 (10%)	44 (7%)
	Past year	4 (12%)	135 (19%)	81 (19%)	94 (19%)	82 (22%)	124 (19%)	123 (19%)
	1-2 years ago	10 (30%)	102 (15%)	49 (11%)	64 (13%)	34 (9%)	58 (9%)	89 (14%)
	2+ years ago	-	214 (31%)	102 (23%)	133 (67%)	103 (27%)	180 (27%)	154 (24%)
NSSI thought(s)								
Lifetime presence		25 (76%)	698 (100%)	418 (96%)	467 (96%)	330 (87%)	633 (95%)	614 (95%)
Lifetime frequency								
	1	0 (0%)	2 (0%)	1 (0%)	-	2 (1%)	3 (1%)	3 (1%)
	2-5	0 (0%)	16 (2%)	17 (4%)	12 (3%)	20 (5%)	8 (1%)	6 (1%)

	5-10	0 (0%)	15 (2%)	9 (2%)	7 (1%)	19 (5%)	12 (2%)	5 (1%)
	10-15	3 (9%)	24 (3%)	8 (2%)	12 (3%)	10 (3%)	16 (2%)	15 (2%)
	15-20	0 (0%)	3 (0%)	11 (3%)	14 (3%)	8 (2%)	13 (2%)	21 (3%)
	20-30	3 (9%)	22 (3%)	13 (3%)	24 (5%)	11 (3%)	18 (3%)	19 (3%)
	30+	18 (55%)	616 (88%)	359 (82%)	398 (82%)	260 (68%)	563 (85%)	545 (84%)
Recency								
	Past month	16 (48%)	537 (77%)	339 (78%)	367 (76%)	239 (63%)	496 (75%)	478 (74%)
	Past year	4 (12%)	99 (14%)	51 (12%)	67 (14%)	59 (16%)	98 (15%)	90 (14%)
	1-2 years ago	4 (12%)	24 (3%)	11 (3%)	16 (3%)	12 (3%)	16 (2%)	17 (3%)
	2+ years ago	-	38 (5%)	17 (4%)	17 (4%)	20 (5%)	23 (4%)	29 (5%)
NSSI Behavior(s)								
Lifetime presence		22 (67%)	678 (97%)	411 (94%)	464 (96%)	321 (85%)	624 (94%)	604 (93%)
Lifetime frequency								
	1	0 (0%)	3 (0%)	2 (1%)	3 (1%)	7 (%)	1 (0%)	6 (1%)
	2-5	0 (0%)	38 (5%)	26 (6%)	25 (5%)	32 (%)	27 (4%)	21 (3%)
	5-10	6 (18%)	30 (4%)	11 (3%)	24 (5%)	24 (%)	24 (4%)	23 (4%)
	10-15	0 (0%)	15 (2%)	28 (6%)	13 (3%)	13 (%)	24 (4%)	30 (5%)
	15-20	0 (0%)	11 (2%)	22 (5%)	24 (5%)	16 (%)	24 (4%)	29 (5%)
	20-30	3 (9%)	29 (4%)	39 (9%)	26 (5%)	14 (%)	38 (6%)	37 (6%)
	30+	24 (73%)	553 (79%)	283 (65%)	349 (72%)	215 (57%)	486 (73%)	458 (70%)
Recency								
	Past month	8 (24%)	356 (51%)	227 (52%)	214 (44%)	135 (36%)	340 (51%)	320 (49%)
	Past year	7 (21%)	181 (26%)	106 (24%)	146 (30%)	112 (30%)	162 (24%)	171 (26%)
	1-2 years ago	7 (21%)	63 (9%)	36 (8%)	43 (9%)	26 (7%)	65 (10%)	45 (7%)
	2+ years ago	-	78 (11%)	42 (10%)	61 (13%)	48 (13%)	57 (9%)	68 (11%)
Co-occurring suicide and NSSI thoughts								
Lifetime presence		10 (29%)	564 (81%)	351 (81%)	384 (79%)	280 (74%)	515 (78%)	497 (77%)
Lifetime frequency								
	1	0 (0%)	13 (2%)	2 (1%)	6 (1%)	6 (2%)	8 (1%)	8 (1%)
	1-5	2 (6%)	85 (10%)	48 (11%)	45 (9%)	49 (13%)	57 (9%)	54 (8%)
	5-10	2 (6%)	27 (4%)	35 (8%)	35 (7%)	30 (8%)	52 (8%)	42 (7%)
	10-15	0 (0%)	13 (2%)	25 (6%)	23 (5%)	23 (6%)	41 (6%)	38 (6%)
	15-20	2 (6%)	5 (1%)	36 (8%)	32 (7%)	19 (5%)	25 (4%)	44 (7%)
	20-30	3 (9%)	10 (1%)	14 (3%)	24 (5%)	19 (5%)	38 (6%)	32 (5%)
	30+	1 (3%)	411 (59%)	191 (44%)	219 (45%)	134 (35%)	294 (44%)	279 (43%)
Recency								
	Past month	3 (9%)	266 (38%)	167 (39%)	177 (37%)	131 (35%)	248 (37%)	247 (38%)
	Past year	4 (12%)	167 (24%)	110 (25%)	104 (21%)	82 (22%)	155 (23%)	147 (23%)
	1-2 years ago	3 (9%)	59 (9%)	37 (9%)	48 (10%)	26 (7%)	54 (8%)	43 (7%)

	2+ years ago	-	72 (10%)	37 (9%)	55 (1%)	41 (11%)	58 (9%)	60 (9%)
Age		31.74 (10.09)	30.52 (10.70)	25.50 (9.73)	26.53 (9.72)	31.18 (11.80)	27.98 (10.10)	28.20 (10.38)
Gender								
	Male	4 (12%)	52 (7%)	28 (6%)	22 (5%)	17 (5%)	30 (5%)	46 (7%)
	Female	15 (45%)	614 (88%)	364 (84%)	399 (82%)	261 (69%)	513 (77%)	506 (78%)
	Transgender/Trans*	-	6 (1%)	4 (1%)	12 (3%)	8 (2%)	25 (4%)	20 (3%)
	Non binary gender	-	20 (3%)	34 (8%)	48 (10%)	10 (3%)	41 (6%)	33 (5%)
	Prefer not to say	-	5 (1%)	2 (1%)	1 (0%)	7 (2%)	3 (1%)	3 (1%)
Ethnicity								
	White	-	-	398 (92%)	447 (92%)	239 (63%)	558 (84%)	554 (85%)
	Mixed/Multiple ethnic groups	-	-	15 (3%)	16 (3%)	15 (4%)	25 (4%)	29 (5%)
	Asian (Indian, Pakistani, Bangladeshi, Chinese)	-	-	6 (2%)	15 (3%)	43 (12%)	15 (2%)	18 (3%)
	Other ethnic group	-	-	9 (2%)	3 (1%)	3 (1%)	10 (2%)	5 (1%)
	Prefer not to say	-	-	4 (1%)	1 (0%)	3 (1%)	4 (1%)	2 (0%)
Marital status								
	Single	10 (30%)	-	286 (66%)	289 (60%)	159 42 (%)	366 (55%)	356 (55%)
	Married	5 (15%)	-	35 (8%)	46 (10%)	47 (12%)	93 (14%)	79 (12%)
	Cohabiting	2 (6%)	-	56 (13%)	82 (17%)	45 (12%)	84 (13%)	89 (14%)
	Separated	1 (3%)	-	14 (3%)	18 (4%)	20 (5%)	14 (2%)	26 (4%)
	Widowed	0 (0%)	-	-	1 (0%)	3 (1%)	2 (0%)	4 (1%)
	Other	1 (3%)	-	38 (9%)	39 (8%)	23 (6%)	45 (7%)	48 (7%)
	Prefer not to say	-	-	3 (1%)	7 (1%)	6 (2%)	8 (1%)	6 (1%)
Employment status								
	Employed	-	-	142 (33%)	171 (35%)	106 (28%)	248 (37%)	243 (37%)
	Self-employed	-	-	18 (4%)	20 (4%)	11 (3%)	22 (3%)	18 (3%)
	Unemployed	-	-	62 (14%)	92 (19%)	68 (18%)	104 (16%)	105 (16%)
	Student	-	-	172 (39%)	157 (32%)	74 (20%)	184 (28%)	191 (29%)
	Retired	-	-	5 (1%)	2 (0%)	7 (2%)	1 (0%)	2 (0%)
	Other	-	-	26 (6%)	32 (7%)	31 (8%)	41 (6%)	40 (6%)
	Prefer not to say	-	-	7 (2%)	8 (2%)	6 (2%)	12 (2%)	9 (1%)
Level of education								
	High/Upper school	-	-	210 (49%)	228 (47%)	108 (29%)	272 (41%)	264 (41%)
	Business/Technical training beyond high/upper school	-	-	68 (16%)	77 (16%)	70 (18%)	126 (19%)	116 (18%)
	Bachelor degree	-	-	103 (24%)	115 (24%)	76 (20%)	138 (21%)	163 (25%)
	Master’s degree	-	-	27 (6%)	31(6%)	28 (7%)	45 (7%)	45 (7%)
	Doctoral degree	-	-	5 (1%)	8 (2%)	5 (1%)	11 (2%)	6 (1%)

	Prefer not to say	-	-	19 (4%)	23 (5%)	16 (4%)	20 (3%)	14 (2%)
Annual income								
	£0-10,000	-	-	241 (55%)	266 (55%)	125 (33%)	299 (45%)	285 (44%)
	£10-20,000	-	-	81 (19%)	92 (19%)	70 (18%)	136 (21%)	138 (21%)
	£20-30,000	-	-	36 (8%)	50 (10%)	30 (8%)	59 (9%)	60 (9%)
	£30-40,000	-	-	13 (3%)	13 (3%)	15 (4%)	30 (5%)	33 (5%)
	£40-50,000	-	-	13 (3%)	5 (1%)	10 (3%)	13 (2%)	18 (3%)
	£50-60,000	-	-	1 (0%)	4 (%)	6 (2%)	6 (1%)	5 (1%)
	£60-70,000	-	-	2 (1%)	-	2 (1%)	3 (1%)	3 (1%)
	£70-80,000	-	-	2 (1%)	-	1 (0%)	3 (1%)	2 (0%)
	£80-90,000	-	-	1 (0%)	-	1 (0%)	3 (1%)	1 (0%)
	£90-100,000	-	-	-	1 (0%)	1 (0%)	1 (0%)	-
	£100,000+	-	-	-	1 (0%)	4 (1%)	2 (0%)	2 (0%)
	Prefer not to say	-	-	42 (10%)	50 (%)	38 (10%)	57 (9%)	61 (9%)

Item Generation and Refinement

A large and varied item pool of beliefs about SIB was generated and refined through several steps (Figure 1). Separate item pools were initially generated by people with lived experience of SIB and the research team. People with lived experience of SIB completed a "thought capture" exercise in which they were asked to report any mental phenomena *about* SIB and its sequalae (thoughts, feelings, mental images, dreams or nightmares, memories, voices, sounds, smells, sensations, and tastes), including cognitions about physical pain, blood, and injuries experienced as part of SIB. Responses to the thought capture exercise were combined to create an item pool. The research team concurrently generated an item pool by drawing on a broad range of sources, using the guiding framework of positive, negative, and facilitating SICs (see Supplementary material for Chapter 4).

The two item pools were combined and then simultaneously refined by clinicians, researchers, and people with lived experience of SIB, retaining any item with unique content. Candidate items were rationally organized into themed groups of items (homogeneous item composites; HICs; Hogan, 1983) to ensure that sufficient markers were included for each potential facet of SICs. The creation of HICs ensured that corresponding factors had a reasonable chance of emerging in our subsequent structural analyses. However, the construction of HICs does not force corresponding factors to emerge.

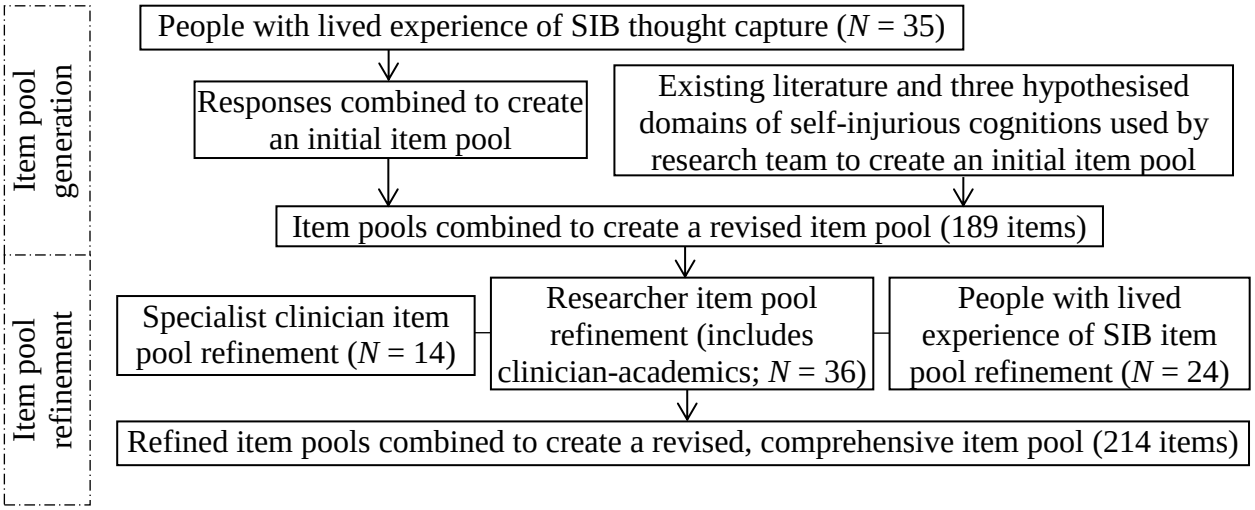


Figure 1. Item generation and refinement steps

4.4. Exploring How Best to Conceptualise Self-Injurious Behavior

Our desire to clarify whether the cognitions that characterise SA are the same as those that characterise NSSI led us to invent a methodology that would directly test this key question. To achieve this, we initially phrased items generically in relation to "self-injurious behavior" and made no reference to suicidal desire or intent[7]. Participants were asked to complete every item twice: once in relation to SA and once in relation to NSSI. We reasoned that different ratings of exactly the same items for the two behaviors would justify the development of separate scales, and that similar ratings of exactly the same items for the two behaviors would justify the development of a single scale.

4.4.1. Measure and Procedure

The item pool consisted of 214 items which were presented in blocks of 25 items per page. The ordering of blocks and items within blocks was randomised. Participants were asked to indicate how much they currently agree with each SIC on a 7-point Likert scale ranging from *Strongly agree* to *Strongly disagree*.

4.4.2. Results

Table 2 presents the results of this analysis. The intraclass correlation coefficient (ICC) is the most informative statistic and examines how identical ratings are to each other by accounting for rank order (whether participants use the Likert scale in the same way) and absolute differences (mean levels). The average ICC across all items is .61, which indicates "moderate" agreement (Fleiss, 1986; Nunnally & Bernstein, 1994). This result provides direct evidence that participants perceived SA and NSSI as similar but not identical constructs.

Separate PAs of the SA and NSSI items (detailed further below) revealed different numbers of factors for SA (18 factors) and NSSI (22 factors) *for exactly the same items*. Overall, these results indicate that participants viewed SA and NSSI as similar but separate constructs. We interpreted this evidence as a rationale to construct separate scales for NSSI and SA. Having established the need to construct separate scales to measure SA and NSSI cognitions, from this point onwards (Samples 2-6) items were rephrased to refer to SA or NSSI (rather than SIB) and separate structural analyses were conducted for SA and NSSI item pools.

[7] The terminology of "self-injurious behaviour" was specifically selected in an attempt to minimize ambiguity and misunderstandings. Several other terms were considered, for example "self-harm," however, these terms tend to carry specific associations of SA or NSSI for members of the general public and may therefore have introduced error.

Table 2 Comparison of Item Responses for Suicide Attempts and Nonsuicidal Self-Injury (Sample 1)

	Suicide attempt		Nonsuicidal self-injury					
Sample 1 Item pool (214 items)	*M*	SD	*M*	SD	Mean difference	Cohen's *d*	*r*	ICC
SIB makes my life better	5.61	1.84	5.19	1.97	0.42	0.22	0.57	0.56
SIB helps me escape from my problems	3.67	2.28	3.35	2.09	0.32	0.15	0.50	0.50
SIB is an escape from myself	3.07	2.21	2.82	1.96	0.25	0.12	0.46	0.46
SIB helps me avoid my problems	4.02	2.33	3.95	2.18	0.07	0.03	0.59	0.59
SIB helps me cope	3.86	2.22	2.20	1.58	1.65	0.86	0.38	0.26
SIB prevents bad things from happening to me	5.38	2.05	5.43	1.89	-0.05	0.03	0.61	0.61
SIB changes the way that I am thinking	3.91	2.11	3.07	1.83	0.83	0.42	0.55	0.50
SIB changes my thoughts so the problems in my life do not seem as bad	4.92	2.06	4.33	2.11	0.60	0.29	0.61	0.59
SIB clears my mind	4.22	2.13	2.94	1.91	1.28	0.63	0.46	0.39
SIB stops negative or distressing thoughts	4.25	2.22	3.20	2.04	1.05	0.49	0.53	0.47
SIB stops upsetting thoughts going round and round in my mind	3.71	2.24	2.78	1.87	0.93	0.45	0.49	0.44
SIB changes images or scenes that run through my mind	4.24	2.11	3.61	2.02	0.62	0.30	0.61	0.59
SIB changes images or scenes that I imagine	4.30	2.11	3.81	2.01	0.48	0.23	0.65	0.63
SIB stops me thinking about upsetting pictures and images	4.30	2.22	3.48	2.13	0.82	0.38	0.63	0.59
SIB interrupts images or scenes that run through my mind	3.99	2.18	3.03	1.90	0.96	0.47	0.58	0.52
SIB interrupts upsetting pictures and scenes that run through my mind	3.91	2.22	3.07	1.98	0.84	0.40	0.63	0.58
SIB stops me feeling numb	4.40	2.13	3.24	2.13	1.16	0.55	0.54	0.47
SIB stops me feeling detached from myself	4.61	2.12	3.44	2.15	1.17	0.55	0.59	0.51
SIB refocuses my mind	4.28	2.13	2.90	1.94	1.38	0.68	0.44	0.36
SIB brings my attention back to my body	4.38	2.13	3.03	1.99	1.34	0.65	0.52	0.43
SIB makes me feel “real” or alive	4.77	2.16	3.33	2.18	1.44	0.66	0.54	0.45
SIB is a way to feel something when I otherwise feel nothing	4.39	2.22	3.13	2.17	1.26	0.57	0.58	0.50
SIB temporarily stops me from feeling anything	3.94	2.23	3.21	2.07	0.73	0.34	0.54	0.51
SIB makes me feel nothing	4.58	2.18	4.40	2.13	0.18	0.08	0.54	0.54

SIB makes me feel detached from myself	3.94	2.23	3.58	2.17	0.36	0.16	0.59	0.58
SIB helps me forget my problems	4.35	2.24	3.60	2.12	0.75	0.34	0.53	0.50
SIB solves my problems	5.30	2.12	5.62	1.80	-0.32	0.16	0.54	0.53
SIB is the best way to change how I feel	4.78	2.09	4.03	2.09	0.75	0.36	0.59	0.55
SIB makes things seem better, even if only for a little while	3.97	2.26	2.62	1.88	1.35	0.65	0.47	0.38
SIB makes me feel less upset	4.27	2.18	2.83	1.90	1.45	0.71	0.43	0.34
SIB is exciting	5.72	1.82	5.40	1.99	0.31	0.16	0.67	0.66
SIB is pleasurable	5.49	1.90	4.30	2.18	1.19	0.59	0.55	0.46
SIB makes me less bored	5.94	1.65	5.69	1.85	0.25	0.14	0.72	0.71
SIB is the best way to calm myself down	4.69	2.11	3.10	1.97	1.59	0.78	0.50	0.38
SIB reduces tension and stress	3.95	2.23	2.22	1.56	1.73	0.90	0.37	0.25
SIB provides relief from upsetting thoughts or feelings	3.35	2.21	2.11	1.54	1.25	0.66	0.42	0.33
SIB is a way to express anger or self-criticism	3.55	2.21	2.27	1.69	1.28	0.65	0.53	0.43
I engage in SIB to punish myself	3.72	2.34	2.67	2.03	1.06	0.49	0.64	0.56
SIB is a form of self-punishment	3.47	2.31	2.43	1.89	1.03	0.49	0.59	0.51
I deserve SIB scars and injuries	3.90	2.29	3.32	2.22	0.59	0.26	0.76	0.73
SIB makes me a better person	6.05	1.52	5.86	1.69	0.19	0.12	0.68	0.67
SIB improves my life	5.82	1.74	5.43	1.91	0.39	0.21	0.57	0.55
SIB makes me who I am	5.25	1.98	4.91	2.11	0.34	0.17	0.72	0.71
SIB is one of the best things in my life	6.20	1.46	6.07	1.62	0.13	0.09	0.69	0.68
I have nothing in my life apart from SIB	5.90	1.68	5.93	1.65	-0.03	0.02	0.77	0.77
SIB is the only thing I've got that’s just mine	5.04	2.15	4.65	2.26	0.40	0.18	0.76	0.74
I am more similar to people who engage in SIB than to people who do not	4.07	1.99	3.80	1.97	0.27	0.14	0.77	0.77
My life would be worse without SIB	5.26	1.96	4.67	2.11	0.59	0.29	0.68	0.65
SIB is a reward or treat for me	6.12	1.54	5.79	1.84	0.33	0.20	0.63	0.61
SIB is comforting	4.23	2.25	2.83	1.91	1.40	0.68	0.51	0.41
SIB keeps me in control	4.26	2.25	3.16	2.04	1.10	0.51	0.54	0.48
SIB makes things more certain, even if only for a while	4.09	2.18	3.68	2.03	0.41	0.20	0.58	0.57
SIB is the only thing that I can control in my life	4.23	2.24	3.85	2.21	0.38	0.17	0.75	0.74
I keep SIB as an option in case my problems get worse	2.97	2.20	2.62	1.91	0.34	0.17	0.58	0.56

Knowing that I can engage in SIB gives me strength to carry on with life	4.29	2.25	3.73	2.18	0.56	0.25	0.64	0.62
SIB is always there to fall back on if things get bad enough	2.90	2.19	2.40	1.85	0.50	0.25	0.59	0.56
SIB releases pressure or tension	3.33	2.20	1.82	1.35	1.51	0.83	0.34	0.23
SIB is a way of maintaining my independence	5.36	1.99	4.99	2.08	0.37	0.19	0.71	0.69
SIB shows that I can cope on my own	5.75	1.78	5.18	2.08	0.56	0.29	0.68	0.64
SIB is a way of caring for myself	5.45	1.96	4.78	2.16	0.67	0.33	0.61	0.58
SIB injuries give me a valid reason to take care of myself	5.07	1.96	4.56	2.09	0.51	0.25	0.69	0.67
SIB injuries give me a legitimate reason to take care of myself	4.99	2.01	4.51	2.12	0.48	0.24	0.71	0.69
I can only be kind towards myself after I have engaged in SIB	5.32	1.91	4.94	2.11	0.38	0.19	0.75	0.73
SIB is easier to deal with than emotional pain	2.87	2.13	1.84	1.41	1.02	0.57	0.45	0.35
Focusing on physical pain takes me mind off my emotional pain	3.07	2.16	1.87	1.43	1.20	0.65	0.47	0.35
SIB injuries remind me of what I have lived through	3.59	2.17	2.83	1.93	0.76	0.37	0.66	0.62
SIB injuries improve my appearance	6.48	1.16	6.51	1.18	-0.04	0.03	0.64	0.65
SIB injuries are comforting	4.65	2.19	3.47	2.19	1.18	0.55	0.57	0.50
SIB injuries show that my feelings are real	3.86	2.20	3.36	2.08	0.50	0.23	0.72	0.70
SIB injuries show that my feelings are important	4.53	2.15	4.35	2.18	0.18	0.08	0.78	0.78
SIB makes me more likeable	6.56	1.02	6.54	1.09	0.01	0.01	0.60	0.60
SIB benefits people around me	5.99	1.77	6.27	1.43	-0.28	0.18	0.58	0.56
SIB protects important people in my life	5.62	1.95	5.42	2.05	0.20	0.10	0.62	0.62
SIB makes people take my problems seriously	4.30	2.24	4.65	2.09	-0.35	0.16	0.73	0.72
SIB makes people take me seriously	4.79	2.19	5.15	1.98	-0.36	0.17	0.74	0.72
SIB makes other people feel guilty	4.26	2.21	4.49	2.10	-0.24	0.11	0.75	0.75
SIB influences other people’s behavior towards me	3.72	2.16	3.68	2.06	0.04	0.02	0.77	0.77
My SIB persuades other people to change their mind	5.93	1.60	6.02	1.55	-0.09	0.06	0.81	0.81
SIB makes other people help me	4.73	2.12	4.95	1.99	-0.22	0.11	0.71	0.70
SIB brings out other people’s true feelings towards me	4.40	2.07	4.43	2.02	-0.03	0.01	0.78	0.78
SIB makes other people understand how distressed I am	3.85	2.24	4.00	2.09	-0.16	0.07	0.73	0.73
SIB helps me connect with other people	6.02	1.56	5.94	1.64	0.08	0.05	0.72	0.72
SIB helps me fit in with other people	6.56	1.04	6.53	1.12	0.03	0.02	0.65	0.65
SIB helps me get accepted by some people	6.22	1.43	6.24	1.43	-0.02	0.01	0.75	0.75
Other people accept me because of SIB	6.28	1.29	6.27	1.32	0.01	0.01	0.77	0.77

SIB makes me feel part of a group	6.38	1.31	6.29	1.44	0.10	0.07	0.67	0.67
Other people approve of my SIB	6.54	1.11	6.53	1.09	0.01	0.01	0.61	0.61
SIB makes people care about me	5.38	1.92	5.49	1.84	-0.10	0.06	0.81	0.81
SIB stops important people in my life from leaving or abandoning me	5.94	1.69	6.05	1.60	-0.11	0.07	0.72	0.72
SIB shows other people how distressed I feel	3.62	2.27	3.69	2.20	-0.07	0.03	0.74	0.74
People who engage in SIB are the only people who understand me	4.50	2.10	4.38	2.11	0.12	0.06	0.86	0.86
SIB helps me escape from harmful relationships	5.02	2.13	5.26	2.04	-0.23	0.11	0.71	0.71
SIB is a form of rebellion	5.85	1.78	5.72	1.88	0.13	0.07	0.74	0.74
SIB shows other people how strong I am	6.16	1.44	6.03	1.60	0.12	0.08	0.69	0.69
SIB is a way to intentionally upset other people	6.18	1.58	6.27	1.48	-0.10	0.06	0.74	0.74
SIB helps me get revenge against others	6.12	1.60	6.22	1.52	-0.10	0.06	0.73	0.73
SIB is a way to get back at people who have hurt me	5.95	1.75	6.07	1.67	-0.12	0.07	0.72	0.72
SIB shows other people that they were wrong	5.76	1.79	5.91	1.71	-0.15	0.09	0.75	0.75
SIB makes people sorry for the way they treated me	5.63	1.90	5.89	1.72	-0.26	0.14	0.76	0.75
SIB prevents me from getting hurt in a worse way	4.46	2.25	4.02	2.24	0.44	0.20	0.51	0.50
SIB prevents me from hurting other people	4.81	2.27	4.53	2.33	0.28	0.12	0.76	0.75
SIB helps me avoid doing things that I do not want to do	5.01	2.16	5.12	2.09	-0.11	0.05	0.63	0.63
SIB stops other people from forcing me to do things	5.46	1.98	5.81	1.71	-0.35	0.19	0.74	0.72
SIB stops people from hurting me	5.25	2.11	5.61	1.89	-0.36	0.18	0.67	0.65
Other people leave me alone because of SIB	4.71	2.05	4.63	2.05	0.08	0.04	0.83	0.83
SIB is the only option I have for solving my problems	4.65	2.15	4.56	2.09	0.09	0.05	0.59	0.59
SIB is the only way to end unbearable pain	3.26	2.30	3.46	2.12	-0.20	0.09	0.52	0.51
SIB is the only method of coping that works for me	4.63	2.12	3.75	2.12	0.88	0.42	0.56	0.52
My problems are so serious that SIB is the only option	4.51	2.22	4.41	2.15	0.09	0.04	0.66	0.66
There are no alternatives to SIB	5.09	2.06	4.90	2.08	0.19	0.09	0.65	0.65
SIB is the only way to control upsetting pictures and images that go through my mind	4.49	2.20	3.81	2.18	0.68	0.31	0.67	0.64
I cannot cope without SIB	4.88	2.06	3.95	2.08	0.93	0.46	0.60	0.54
I rely on SIB	4.94	2.06	3.86	2.12	1.08	0.52	0.51	0.45
I cannot function without SIB	5.25	1.90	4.48	2.10	0.77	0.39	0.63	0.59

If I stopped SIB, I would be overwhelmed by distressing thoughts and feelings	4.14	2.21	3.19	2.08	0.95	0.44	0.60	0.55
The distress that I would feel if I stopped SIB would be unbearable	4.65	2.10	3.91	2.14	0.74	0.36	0.62	0.58
I would lose control without SIB	4.72	2.10	3.85	2.15	0.88	0.41	0.62	0.57
I need to always have SIB as an option in my life	3.64	2.30	3.12	2.10	0.53	0.24	0.61	0.59
SIB is the best way to deal with my problems	5.20	2.05	4.87	2.12	0.33	0.16	0.60	0.60
SIB is better than the alternative options	4.54	2.25	3.71	2.16	0.82	0.38	0.49	0.46
SIB creates a lot of problems for me	3.30	2.08	3.20	1.92	0.10	0.05	0.59	0.59
SIB makes my problems worse	3.68	2.04	3.96	1.94	-0.28	0.15	0.57	0.56
SIB has ruined my life	4.67	2.06	4.69	2.05	-0.02	0.01	0.73	0.73
SIB makes me more distressed than if I did not engage in SIB	4.26	2.16	4.77	1.99	-0.51	0.25	0.59	0.57
SIB proves that I am impulsive	4.07	2.25	3.72	2.20	0.35	0.16	0.77	0.76
SIB has made me crazy	4.69	2.15	4.68	2.11	0.01	0.00	0.81	0.81
SIB is abnormal or bad	3.79	2.19	3.73	2.12	0.06	0.03	0.82	0.82
The fact that I engage in SIB makes me a bad person	4.53	2.28	4.49	2.30	0.04	0.02	0.82	0.82
My life would be better without SIB	2.86	2.00	2.89	1.89	-0.03	0.02	0.63	0.63
SIB is wrong	3.94	2.21	3.99	2.15	-0.05	0.02	0.78	0.78
SIB is a private act	1.91	1.69	1.37	0.97	0.54	0.39	0.36	0.29
SIB is destructive	2.20	1.81	2.23	1.64	-0.03	0.02	0.54	0.54
SIB is physically painful	2.96	1.99	2.51	1.73	0.45	0.24	0.55	0.53
SIB is a bad option	2.88	2.03	3.09	1.95	-0.21	0.11	0.60	0.60
SIB brings my most important beliefs into question	3.99	2.18	4.30	2.05	-0.31	0.15	0.78	0.77
Alternatives to SIB are better than SIB	3.19	2.08	3.41	2.00	-0.22	0.11	0.62	0.62
There are alternatives to SIB	2.95	2.01	2.80	1.82	0.14	0.07	0.66	0.66
There are solutions to my problems other than SIB	3.09	2.06	2.92	1.89	0.17	0.09	0.65	0.64
I hate taking care of my SIB injuries	3.82	2.03	3.82	2.09	0.00	0.00	0.69	0.69
I hate my SIB injuries	3.61	2.09	3.38	2.13	0.23	0.11	0.68	0.68
SIB scars or injuries are unattractive	3.13	2.09	2.75	1.98	0.38	0.19	0.75	0.74
SIB makes me different to other people	3.15	1.97	2.97	1.82	0.18	0.10	0.77	0.77
SIB damages important relationships in my life	2.96	2.06	2.97	1.94	0.00	0.00	0.68	0.68
SIB upsets or hurts the people I care about	2.25	1.84	2.17	1.61	0.07	0.04	0.60	0.59

SIB makes me a social outcast	3.68	2.06	3.55	2.01	0.13	0.07	0.81	0.81
I have to hide my SIB	2.23	1.78	1.84	1.43	0.40	0.25	0.44	0.42
I think less of someone when I learn that they engage in SIB	6.29	1.38	6.22	1.52	0.07	0.05	0.70	0.70
SIB creates a barrier between myself and others	2.94	1.93	2.75	1.75	0.19	0.11	0.71	0.70
SIB creates a boundary between myself and others	3.05	1.96	2.80	1.77	0.25	0.13	0.71	0.70
I cannot talk to other people about my SIB	2.48	1.83	2.35	1.61	0.13	0.08	0.64	0.63
People exclude me because of my SIB	4.24	2.05	4.18	2.03	0.06	0.03	0.83	0.83
People reject me because of my SIB	3.85	2.05	3.77	2.00	0.08	0.04	0.82	0.82
People punish my SIB	4.25	2.11	4.13	2.08	0.12	0.06	0.75	0.75
People think that my SIB is selfish	2.64	1.88	2.80	1.79	-0.16	0.09	0.65	0.65
People think less of me when they discover my SIB	3.13	1.91	2.85	1.74	0.28	0.15	0.76	0.74
People do not understand my SIB	2.23	1.67	1.94	1.32	0.29	0.20	0.55	0.53
People try to stop my SIB	2.79	2.03	2.87	1.94	-0.08	0.04	0.64	0.64
People judge and criticise my SIB	2.88	1.91	2.65	1.72	0.22	0.13	0.74	0.73
SIB leads to unwanted attention from other people	3.04	1.99	2.70	1.81	0.34	0.18	0.63	0.62
Other people are not interested in my SIB	3.57	1.89	3.27	1.73	0.29	0.17	0.75	0.74
People think that my SIB is bad or wrong	2.27	1.74	2.00	1.39	0.26	0.17	0.61	0.59
People think that my SIB is abnormal	2.56	1.78	2.25	1.52	0.31	0.19	0.66	0.64
SIB controls me	5.24	1.99	4.76	2.10	0.48	0.23	0.65	0.63
SIB controls my life	5.37	1.94	5.00	2.05	0.38	0.19	0.68	0.66
I need to regularly engage in SIB	5.56	1.79	4.43	2.12	1.13	0.58	0.51	0.43
SIB is an addiction	4.54	2.21	3.03	2.12	1.52	0.72	0.52	0.42
I am addicted to SIB	5.47	1.93	4.21	2.29	1.27	0.60	0.53	0.44
My SIB will get worse	4.44	2.07	4.23	2.02	0.21	0.11	0.75	0.74
SIB has taken over my life	5.20	2.03	4.72	2.14	0.49	0.23	0.62	0.61
I cannot control my SIB	4.80	2.09	4.30	2.08	0.50	0.25	0.64	0.63
I cannot stop SIB	4.87	2.06	4.08	2.07	0.78	0.38	0.58	0.54
I am scared that I will injure myself more severely than I intend	4.27	2.34	3.54	2.25	0.73	0.33	0.56	0.53
I have no control over how I injure myself	5.48	1.86	5.31	1.89	0.17	0.09	0.70	0.70
I have very little influence over my SIB	4.82	2.04	4.50	2.02	0.33	0.17	0.68	0.68
Thoughts and urges to engage in SIB are overwhelming	3.05	2.11	2.49	1.75	0.56	0.29	0.58	0.55
I feel on edge if I don't engage in SIB	4.86	2.05	3.61	2.11	1.25	0.62	0.54	0.46

I feel compelled to engage in SIB if I notice that my injuries are beginning to heal	5.40	1.92	4.33	2.33	1.07	0.50	0.62	0.54
Urges to engage in SIB cannot be resisted	4.60	2.07	3.94	2.06	0.66	0.33	0.62	0.59
I'm not strong enough to stop SIB	4.51	2.15	3.98	2.15	0.53	0.25	0.67	0.65
I give myself permission to engage in SIB	4.27	2.21	3.29	2.08	0.98	0.47	0.54	0.49
I vow that 'This will be the last time I engage in SIB'	4.22	2.40	4.43	2.32	-0.21	0.09	0.65	0.65
I think things that sabotage my efforts to avoid SIB	4.09	2.02	3.78	2.01	0.30	0.15	0.68	0.67
I think things that make SIB more likely to happen	3.82	2.10	3.51	2.09	0.32	0.15	0.75	0.75
I have good reasons for SIB	3.33	2.19	2.94	1.99	0.39	0.19	0.68	0.66
I encourage myself to engage in SIB	5.41	1.94	5.10	2.02	0.31	0.16	0.64	0.64
I tell myself that I can engage in SIB if I do something to make up for it afterwards	6.00	1.56	5.67	1.79	0.33	0.20	0.68	0.66
It's acceptable to engage in SIB if I do it in a particular way	5.00	2.04	4.17	2.15	0.83	0.40	0.59	0.55
SIB is more acceptable if I have opportunity to do it	5.01	1.95	4.61	2.01	0.40	0.21	0.68	0.67
I try not to think about the disadvantages of SIB	3.99	2.25	3.48	2.11	0.51	0.23	0.63	0.61
I ignore problems associated with SIB	4.01	2.12	3.56	2.04	0.45	0.22	0.70	0.69
I under-estimate the consequences of SIB	3.88	2.24	3.56	2.08	0.32	0.15	0.60	0.59
I try to ignore the physical pain that comes with SIB	3.99	2.27	4.10	2.34	-0.10	0.05	0.64	0.64
I tell myself that I deserve the benefits of SIB	4.52	2.21	4.29	2.22	0.23	0.10	0.68	0.68
I find ways to justify SIB to myself	3.71	2.24	3.22	2.13	0.49	0.23	0.63	0.61
It's acceptable to engage in SIB if I'm really upset	4.45	2.17	3.49	2.11	0.96	0.45	0.62	0.57
The benefits of SIB are worth the risks	4.48	2.26	3.53	2.09	0.95	0.45	0.51	0.47
I think that SIB won't be dangerous as long as I am careful	5.11	2.14	3.10	2.10	2.01	0.95	0.37	0.26
SIB is a problem for some people but it won't be for me	5.02	1.98	4.75	2.08	0.26	0.13	0.68	0.67
SIB is more acceptable if I do something to make up for it afterwards	5.82	1.61	5.61	1.73	0.21	0.13	0.73	0.73
I allow myself to delay when I'll engage in SIB, knowing that I will do it later	4.15	2.12	3.74	2.10	0.41	0.20	0.50	0.50
Average values	4.44	2.02	4.02	1.94	0.42	0.24	0.64	0.61

Note: ICC = Single Measure Two-Way Mixed Absolute Agreement Intraclass Correlation Coefficient; Shrout & Fleiss (1979); SIB = self-injurious behavior.

4.5. Preliminary Item Pool Analyses: Sample 1

The goal at this point was to evaluate the nature and quality of the items in our two initial item pools and to identify potentially interesting SICs that might be underrepresented.

4.5.1. Results

We began by removing one item ("SIB is comforting") that had inadvertently been included in the item pool twice. 12 strongly correlated items were deleted from the NSSI item pool and 3 strongly correlated items were deleted from the SA item pool. PAs were then conducted and indicated 22 factors for NSSI and 18 factors for SA. At this stage, the 11 *Anti-suicide* items were analysed separately because, in contrast to the rest of the item pool, these items had referred to "NSSI" rather than "SIB." A PA indicated 3 factors. We note, however, that PAs of NSSI or SA items which included the *Anti-suicide* items produced very similar, although less interpretable, factor structures.

Cognitions about Attempting Suicide. An EFA suggested the existence of 11 potentially meaningful factors. These were tentatively labelled *Dependence* (26 items; e.g., "I would lose control without SIB"), *Escape* (23 items; e.g., "SIB stops negative or distressing thoughts"), *Stigma* (22 items; e.g., "People judge and criticise my SIB"), *Problematic* (20 items; e.g., "SIB is wrong"), *Belonging* (10 items; e.g., "SIB helps me fit in with other people"), *Eliciting help* (11 items; e.g., "SIB makes people take my problems seriously"), *Revenge* (11 items; e.g., "SIB is a way to get back at people who have hurt me"), *Self-validation* (14 items; e.g., "SIB injuries give me a legitimate reason to take care of myself"), *Valued behavior* (15 items; e.g., "SIB makes me a better person"), and *Facilitating cognitions* (11 items; e.g., "SIB is more acceptable if I do something to make up for it afterwards"). One factor, tentatively labelled *Self-punishment* (4 items; e.g., "I engage in SIB to punish myself"), appeared to be underrepresented in our initial SA item pool. That is, several items correlated together reasonably strongly to form what appeared to be a meaningful theme. Five additional items were generated (forming an 8-item HIC) to explore whether doing so would allow a meaningful factor to emerge.

Cognitions about Nonsuicidal Self-Injury. An EFA suggested the existence of 16 potentially meaningful factors. These were very tentatively labelled *Dependence* (28 items; e.g., "I cannot cope without SIB"), *Release* (13 items; e.g., "SIB provides relief from upsetting thoughts or feelings"), *Stigma* (17 items; e.g., "People judge and criticise my SIB"), *Problematic* (22 items; e.g., "I hate my SIB injuries"), *Eliciting help* (10 items; e.g., "SIB makes people take my problems seriously"), *Imagery change* (5 items; e.g., "SIB interrupts images or scenes that run through my mind"), *Revenge* (11 items; e.g., "SIB helps me get

revenge against others"), *Facilitating cognitions* (15 items; e.g., "SIB is more acceptable if I do something to make up for it afterwards"), *Self-validation* (11 items; e.g., "SIB injuries give me a legitimate reason to take care of myself"), *Belonging* (7 items; e.g., "SIB helps me fit in with other people").

Five factors appeared to be underrepresented in our initial NSSI item pool. These were tentatively labelled, *Enjoyable* (6 items; e.g., "SIB is exciting"), *Anti-dissociation* (5 items; e.g., "SIB stops me feeling numb"), *Escape* (6 items; e.g., "SIB helps me escape from my problems"), *Backup plan* (6 items; e.g., "I keep SIB as an option in case my problems get worse"), and *Self-punishment* (5 items; e.g., "I engage in SIB to punish myself"). Additional items were generated for each of these potential factors to explore whether doing so would allow corresponding factors to emerge.

4.5.2. Discussion

These results demonstrated that it is possible to identify specific, differentiable components of SICs. However, an obvious limitation arises from the possibility that the structural analyses may have been confounded by the design or by the administration of a large item pool (214 items were completed twice). To address this, we dropped the two-column methodology, rephrased items to refer to SA or NSSI, and administered fewer items.

4.6. Preliminary Item Pool Analyses: Sample 2

This study aimed to explore the generalisability of the factor structure obtained using Sample 1 and to examine whether adding new items encouraged additional factors to clearly emerge. In total, 62 SA items and 118 NSSI items (including all the Anti-suicide items) were administered.

4.6.1. Results

PAs indicated 8 SA factors and 13 NSSI factors.

Cognitions about Attempting Suicide. An EFA suggested the existence of 6 potentially meaningful factors, which were tentatively labelled *Belonging* (e.g., "Attempting suicide helps me fit in with other people"), *Dependence* (e.g., "Attempting suicide is the only option I have for solving my problems"), *Interpersonal influence* (e.g., "Attempting suicide makes people take my problems seriously"), *Self-punishment* (e.g., "Attempting suicide is an expression of my self-hatred"), *Problematic* (e.g., "Attempting suicide makes my problems worse") and *Escape* (e.g., "Attempting suicide changes the way that I am thinking").

Cognitions about Nonsuicidal Self-Injury. An EFA suggested the existence of 12 potentially meaningful factors, which were tentatively labelled *Escape* (e.g., "NSSI stops my emotional pain"), *Anti-suicide* (e.g., "NSSI is a compromise instead of killing myself"),

Enjoyable (e.g., "NSSI is enjoyable"), *Interpersonal influence* (e.g., "NSSI makes people take my problems seriously"), *Backup plan* (e.g., "I keep NSSI as an option in case my problems get worse"), *Anti-dissociation* (e.g., "NSSI shocks my body so I begin feeling again"), *Self-punishment* (e.g., "I engage in NSSI because I deserve to suffer"), *Self-validation* (e.g., "NSSI makes me a better person"), *Problematic* (e.g., "NSSI makes my problems worse"), *Stigma* (e.g., "People judge and criticise my NSSI"), *Imagery change* (e.g., "NSSI interrupts images or scenes that run through my mind"),and *Belonging* (e.g., "NSSI helps me fit in with other people").

4.6.2. Discussion

The factor structure for SA and NSSI observed in Sample 1 generally replicated in Sample 2, although fewer factors emerged and several of the factors were not particularly clearly defined (having no obvious single theme). Taken together, the structural results observed in Samples 1 and 2 did not provide a definitive indication of the factor structure of cognitions about SA or cognitions about NSSI. We addressed this issue by collecting additional independent samples and administering most of the items that had been administered to Sample 1 as well as the items that had been generated for Sample 2, ensuring that multiple markers (at least 6 items so as to form a HIC; Hogan, 1983) were included for all potential factors.

4.7. Development of the Final Scales: Samples 3, 4, and 5

Three additional samples were collected and each was subjected to a separate EFA (see Tables 3 and 4 and Supplementary material for Chapter 4). The same factor structure emerged across Samples 3, 4, and 5, indicating that robust and replicable factor structures characterize SA and NSSI cognitions. Table 5 presents sample items and summarizes the theme conveyed by each subscale that formed the final version of the SABS and NSIBS.

Cognitions about Attempting Suicide. 7 clearly interpretable and differentiated factors emerged across Samples 3-5 to characterize SA cognitions. These were labelled: *Belonging*, *Stigma*, *Self-punishment*, *Eliciting help*, *Escape*, *Dependence*, and *Revenge*. Comparing the final factor structure (Table 3) to that obtained in the early stages of scale development, six of the seven SABS factors were apparent in Sample 1 (when items has referred to "SIB"). *Self-punishment* had few items and was only a potential factor at that stage. Four of the seven SABS factors were apparent in Sample 2 (*Belonging, Self-punishment, Escape, Dependence*). The *Stigma* factor was not apparent in Sample 2 and the *Eliciting help* and *Revenge* factors merged to form an *Interpersonal influence* factor.

Cognitions about Nonsuicidal Self-Injury. 10 clearly interpretable and differentiated factors emerged across Samples 3-5 to characterize NSSI cognitions. These were labelled: *Escape*, *Self-punishment*, *Anti-dissociation*, *Interpersonal influence*, *Stigma*, *Dependence*, *Problematic*, *Anti-suicide*, *Enjoyable*, and *Belonging*. One caveat to the NSSI structural results is that in Sample 4, the *Interpersonal influence* and *Belonging* factors merged, presumably because the items were so closely themed/correlated. Comparing the final factor structure (Table 4) to that obtained in the early stages of scale development, five of the ten SABS factors (*Dependence*, *Stigma*, *Problematic*, *Belonging*, *Enjoyable*) were apparent in Sample 1 (when items has referred to "SIB"). Nine of the ten NSIBS factors were apparent in Sample 2 (*Dependence* was not apparent).

Retaining Items for the Finished SABS and NSIBS. To enhance the usability of the SABS and NSIBS in clinical practice (e.g., individuals who engage in SIB are often very distressed and often have impaired cognition and comorbid psychological problems), we only retained a subset of items from each factor. We retained a minimum of 3 items per factor (Tabachnick & Fidell, 2007) and selected items that tapped different facets of each construct, ensuring that each subscale had at least a good level of internal consistency ($\alpha \geq .8$; Clark & Watson, 1995; Nunnally 1978; see Part II). We expected that a briefer measure would reduce fatigue and response biases and be most appropriate for this population. Tables 3 and 4 present the final obtained factor structures of the SABS and NSIBS; the underlined items denote the items that we retained and which are used in the Part II analyses.

Facilitating Cognitions. Items designed to measure facilitating cognitions were administered to Samples 1-3 in an attempt to encourage a commensurate factor to emerge. However, a commensurate factor did not emerge for either SA or NSSI across our structural analyses. The items designed to measure facilitating cognitions loaded onto a variety of factors and often cross-loaded. These items were therefore omitted from Samples 4 and 5 and do not feature in the final versions of the SABS or NSIBS[8].

[8] During the development of the SABS and NSIBS, we also explored whether it would be informative to include items tapping how SA and NSSI impair functioning or cause clinical significant distress. However, these items performed poorly in our structural analyses. They were omitted from the SABS and NSIBS (see Supplementary material for Chapter 5).

Table 3 Promax-rotated Loadings of Cognitions about Attempting Suicide in Sample 5

	Factor loadings						
	1	2	3	4	5	6	7
1. *Belonging*							
Attempting suicide helps me fit in with other people	**.850**	.038	.042	-.085	-.082	.059	.073
Attempting suicide helps me get accepted by some people	**.848**	.052	.026	.110	-.018	-.048	-.070
Attempting suicide makes me feel part of a group	**.794**	-.011	.042	-.027	-.085	.060	.048
Attempting suicide helps me connect with other people	**.771**	.015	-.048	.125	.061	.003	-.042
Other people accept me because of my suicide attempt(s)	**.701**	.028	.036	-.026	.121	-.083	.007
2. *Stigma*							
People think that my suicide attempt(s) are selfish	.019	**.863**	-.057	.007	.012	.038	.008
People judge and criticise my suicide attempt(s)	.055	**.800**	.009	-.034	-.002	.106	-.034
People think that my suicide attempt(s) are abnormal	.007	**.789**	-.040	.043	.036	-.040	.002
People punish my suicide attempt(s)	.076	**.710**	.015	-.159	-.093	.119	.173
Attempting suicide damages important relationships in my life	-.059	**.688**	.091	.021	.022	-.114	.083
Attempting suicide leads to unwanted attention from other people	.034	**.595**	.016	.099	.090	-.059	-.127
3. *Self-punishment*							
I attempt suicide because I deserve to suffer	.006	-.021	**1.010**	-.038	-.015	-.038	-.045
I attempt suicide to punish myself	.008	.034	**.925**	-.152	.056	-.111	.096
I deserve suicide attempt scars and injuries	.125	.055	**.800**	.002	-.045	.012	-.162
I attempt suicide to show how much I hate myself	-.004	-.068	**.728**	.135	.036	.031	.020
I attempt suicide because I am worthless and unlovable	-.073	.168	**.566**	.141	-.092	.210	-.080
Attempting suicide is a way to express anger or self-criticism	-.100	.014	**.402**	.067	.180	-.035	.319
4. *Eliciting help*							
Attempting suicide makes people take my problems seriously	.003	.003	-.125	**.899**	.050	-.023	-.025
Attempting suicide makes other people help me	.077	-.043	.009	**.877**	-.071	.020	-.079
Attempting suicide makes other people understand how distressed I am	-.060	.045	.021	**.742**	.005	-.034	.124
Attempting suicide shows other people how distressed I feel	-.057	.115	.112	**.697**	-.040	.030	.024
Attempting suicide makes people care about me	.199	-.092	.012	**.609**	.015	-.042	.101
5. *Escape*							
Attempting suicide changes the way that I am thinking	.039	.094	-.020	-.078	**.885**	-.229	-.010
Attempting suicide clears my mind	.061	-.031	.048	-.087	**.775**	.061	-.048
Attempting suicide changes my thoughts so the problems in my life do not seem as bad	.001	-.038	-.066	.078	**.762**	-.059	.078
Attempting suicide helps me forget my problems	.031	-.089	.056	.061	**.666**	.138	-.064
Attempting suicide stops upsetting thoughts going round and round in my mind	-.145	.104	-.039	.069	**.607**	.193	-.085
Attempting suicide temporarily stops me from feeling anything	-.010	.182	.127	.016	**.484**	.087	-.112
6. *Dependence*							
Attempting suicide is the only option I have for solving my problems	-.099	.110	-.024	.091	-.138	**.830**	-.068

Attempting suicide is the only method of coping that works for me	.050	.020	-.030	-.069	.043	**.786**	-.068
My life would be worse without suicide attempts	.252	-.150	.118	-.042	.039	**.447**	.035
Attempting suicide makes my life better	.263	-.092	-.126	-.035	.219	**.432**	.081
Attempting suicide is the only thing I've got that's just mine	-.001	-.042	.068	-.101	.124	**.431**	.292
Attempting suicide is the only way to control upsetting pictures and images that go through my mind	-.114	.069	.089	-.009	.301	**.428**	.000
7. *Revenge*							
Attempting suicide is a way to get back at people who have hurt me	-.020	-.031	.021	.025	-.114	.036	**.874**
Attempting suicide is a way to intentionally upset other people	.147	.098	-.096	-.018	-.027	-.097	**.674**
Attempting suicide shows other people that they were wrong	.075	-.021	-.035	.176	-.038	.125	**.580**
Attempting suicide is a form of rebellion	.188	.034	-.031	.042	.109	-.149	**.514**
Attempting suicide stops other people from forcing me to do things	.074	.022	.013	.205	.117	.066	**.287**

Note. The highest factor loading for each item is highlighted. Underlined items form the final version of the Suicide Attempt Beliefs Scale (SABS)

Table 4 Promax-rotated Loadings of Cognitions about Nonsuicidal Self-Injury (NSSI) in Sample 5

	Factor loadings									
	1	2	3	4	5	6	7	8	9	10
1. *Escape*										
NSSI helps me escape negative emotions	**.856**	-.053	-.003	-.068	-.073	.012	.038	.045	-.025	.029
NSSI provides relief from upsetting thoughts or feelings	**.828**	.005	-.014	-.101	.055	-.016	.028	-.033	-.054	.045
NSSI makes me feel less upset	**.655**	-.090	.051	.024	.024	-.003	-.060	.030	.116	-.005
NSSI reduces tension and stress	**.618**	-.061	-.067	.044	-.080	.056	.027	.026	.295	-.089
NSSI helps me escape feeling defeated or helpless	**.577**	.061	.106	.051	-.014	-.057	-.009	.047	.039	.028
NSSI helps me escape from my problems	**.567**	.037	.014	.046	.010	.159	.035	-.009	.059	-.001
NSSI temporarily stops me from feeling anything	**.464**	.049	-.067	-.054	.096	.060	.083	-.024	-.007	.154
NSSI stops me thinking about upsetting pictures and images	**.442**	.095	.078	.073	.174	.086	-.093	-.058	-.113	-.012
2. *Self-punishment*										
I engage in NSSI because I deserve to suffer	-.107	**.848**	.005	-.058	.041	.092	-.041	.009	.004	.026
I use NSSI to punish myself	-.021	**.839**	.031	-.016	-.032	.019	.041	-.053	-.021	-.020
NSSI is an expression of my self-hatred	.078	**.836**	-.030	.039	-.046	-.072	.076	-.013	-.025	.006
I engage in NSSI to show how much I hate myself	-.043	**.834**	-.022	.088	-.027	-.071	.018	.031	.050	.006
I engage in NSSI because I am worthless and unlovable	.009	**.726**	.016	-.034	.084	.070	-.061	.046	-.071	-.018
I deserve NSSI scars and injuries	-.120	**.614**	.002	-.146	.033	.207	-.086	.063	.110	.113
NSSI is a way to express anger or self-criticism	.258	**.567**	.017	.136	-.055	-.095	.130	-.059	.026	-.121
3. *Anti-dissociation*										
NSSI stops me feeling numb	.001	.053	**.913**	-.020	-.056	-.020	.000	-.057	-.018	.006
NSSI is a way to feel something when I otherwise feel nothing	-.079	.076	**.845**	-.059	-.002	-.042	-.034	.069	.015	.034
NSSI stops me feeling detached from myself	.048	-.054	**.826**	.025	-.016	.007	.007	.006	-.101	.025
NSSI takes me out of a detached state	.054	-.105	**.818**	.039	-.065	.099	.064	-.005	-.055	-.061
NSSI shocks my body so I begin feeling again	.026	.017	**.757**	.037	.045	-.022	-.002	-.007	.011	-.008
NSSI makes me feel “real” or alive	.014	.042	**.634**	-.025	.037	-.054	-.040	.023	.185	-.016
4. *Interpersonal influence*										
NSSI is a way to intentionally upset other people	-.066	-.007	-.117	**.720**	.066	.016	-.031	.019	.010	-.023
NSSI makes people sorry for the way they treated me	.007	-.051	-.003	**.719**	.019	.005	-.005	.035	-.004	.034
NSSI makes people care about me	-.002	.032	.047	**.711**	-.023	-.015	-.039	.030	.024	.033
NSSI makes people take my problems seriously	.004	.054	.120	**.691**	.018	-.030	.009	-.005	-.033	-.104
My NSSI persuades other people to change their mind	.013	-.021	-.048	**.649**	-.008	.042	.057	-.043	.015	.136
NSSI stops important people in my life from leaving or abandoning me	-.047	.013	.009	**.647**	-.016	.155	.028	-.042	-.006	.046
5. *Stigma*										
People reject me because of my NSSI	-.058	-.012	-.026	.044	**.833**	.063	-.006	-.065	-.076	.012
People judge and criticise my NSSI	.032	-.042	-.052	.020	**.833**	-.096	.022	.064	.034	-.035
People punish my NSSI	-.030	.038	-.008	.049	**.787**	.036	-.117	-.090	-.058	.040

People think that my NSSI is selfish	.021	-.001	.016	-.017	**.787**	-.043	.014	.030	-.047	.038
People think that my NSSI is abnormal	.127	-.019	-.029	-.026	**.587**	-.073	.122	.030	.081	-.090
People do not understand my NSSI	.143	.117	.059	-.081	**.462**	-.022	.049	.074	.031	-.054
6. *Dependence*										
I cannot cope without NSSI	.034	-.002	-.013	-.001	-.038	**.909**	-.058	-.029	-.042	-.006
NSSI is the only method of coping that works for me	.156	-.046	.062	.018	-.014	**.772**	-.077	.006	-.054	-.045
My NSSI will get worse	-.084	.086	-.011	.024	-.054	**.704**	.107	.013	-.025	.025
My problems are so serious that NSSI is the only option	.131	-.038	.008	.111	.074	**.590**	-.047	.091	-.088	-.021
I need to always have NSSI as an option in my life	.045	.071	-.042	.032	.004	**.548**	-.147	.085	.163	-.059
I feel on edge if I don't engage in NSSI	.156	.091	-.065	-.012	.024	**.415**	.044	-.046	.236	-.018
7. *Problematic*										
NSSI makes my problems worse	-.200	.005	.008	.041	.002	-.071	**.772**	.022	.161	-.012
My life would be better without NSSI	.197	.051	-.041	.002	-.059	-.269	**.715**	.011	-.142	.019
NSSI is destructive	.074	.021	-.022	.049	.033	-.123	**.601**	.054	.019	-.021
NSSI creates a lot of problems for me	-.155	-.072	.033	-.030	.236	.171	**.584**	-.032	.137	-.034
I hate my NSSI injuries	.151	.043	.057	-.054	-.107	.167	**.563**	-.036	-.276	.028
NSSI has ruined my life	-.090	-.025	.013	-.043	.048	.410	**.547**	-.014	-.056	.060
8. *Anti-suicide*										
NSSI is a compromise instead of killing myself	-.032	-.011	-.042	.006	-.056	.058	.074	**.888**	.013	-.007
NSSI is a replacement for suicidal behavior	.015	.019	.019	-.044	-.051	-.004	-.006	**.861**	-.013	.035
I deliberately use NSSI to avoid acting on suicidal thoughts	.002	-.015	.052	-.044	.093	.060	-.040	**.794**	-.067	.028
NSSI lets me express my suicidal thoughts without risking death	.066	.036	-.005	.100	-.009	-.043	.001	**.745**	.010	-.036
9. *Enjoyable*										
NSSI is enjoyable	.057	.037	-.106	.010	-.028	-.059	-.060	.021	**.808**	-.060
NSSI is satisfying	.285	.043	-.037	-.059	-.090	.005	.042	-.014	**.640**	-.048
NSSI is uplifting	.224	-.056	.002	-.047	-.031	.078	-.102	-.082	**.564**	.101
NSSI leaves me feeling energized	.101	-.015	.176	.045	.058	-.052	-.076	-.050	**.504**	.058
NSSI makes me less bored	-.104	.017	.031	.210	-.049	-.030	.063	.030	**.488**	.111
NSSI gives me a “high” that feels like a drug high	.028	-.062	.213	-.071	.120	.037	.105	-.012	**.453**	.022
10. *Belonging*										
NSSI helps me fit in with other people	.011	.031	.001	.030	-.006	-.013	-.033	.008	-.006	**.841**
NSSI helps me get accepted by some people	.040	.053	-.036	.116	.025	-.081	.019	-.008	.022	**.751**
NSSI makes me more likeable	.034	-.074	.000	.079	-.089	.049	.041	.068	.060	**.640**
Other people accept me because of NSSI	.003	.020	.009	.110	-.020	.014	.001	-.045	-.043	**.620**
NSSI helps me connect with other people	.038	-.051	.026	.190	.100	-.062	-.009	.010	.005	**.594**

Note. The highest factor loading for each item is highlighted. Underlined items form the final version of the Nonsuicidal Self-Injury Beliefs Scale (NSIBS).

Table 5 Summary of the Themes Conveyed by the Suicide Attempt Beliefs Scale (SABS) and Nonsuicidal Self-Injury Beliefs Scale (NSIBS) Subscales

Scale/Subscale	Sample items	Perceptions about how SA or NSSI relate to oneself or other people	Description
	SABS		
Self-punishment	I attempt suicide because I deserve to suffer I attempt suicide because I am worthless and unlovable	Individual	Perceptions that attempting suicide is a means to punish oneself for being fundamentally flawed, bad, and unlovable
Escape	Attempting suicide stops upsetting thoughts going round and round in my mind Attempting suicide temporarily stops me from feeling anything	Individual	Perceptions that attempting suicide provides relief, for a while, from intolerable cognitive and emotional experiences (perceptions, perseverative thoughts, distress)
Dependence	Attempting suicide is the only option I have for solving my problems Attempting suicide is the only thing I've got that's just mine	Individual	Perceptions that attempting suicide is the only option for self-regulation; that there are no alternative options; that attempting suicide provides important advantages; and a (counterintuitive) perception that suicide attempts are not necessarily/always expected to result in death
Belonging	Attempting suicide helps me fit in with other people Attempting suicide helps me get accepted by some people	Interpersonal	Perceptions that attempting suicide leads to meaningful emotional connections and support from other people
Stigma	People think that my suicide attempt(s) are selfish Attempting suicide damages important relationships in my life	Interpersonal	Perceptions that other people do not respond empathically to suicide attempts and instead judge, criticise, or punish these behaviors
Eliciting help	Attempting suicide makes other people help me Attempting suicide shows other people how distressed I feel	Interpersonal	Perceptions that attempting suicide forces or powerfully elicits empathy, validation, and meaningful support from other people. The motivation is to elicit help from others

Revenge	Attempting suicide is a way to intentionally upset other people Attempting suicide shows other people that they were wrong NSIBS	Interpersonal	Perceptions that attempting suicide is a means to rebel, prove others wrong, or obtain revenge for mistreatment. The motivation is to emotionally hurt others
Self-punishment	I engage in NSSI because I deserve to suffer I engage in NSSI because I am worthless and unlovable	Individual	Perceptions that NSSI is a means to punish oneself for being fundamentally flawed, bad, and unlovable
Escape	NSSI helps me escape negative emotions NSSI helps me escape from my problems	Individual	Perceptions that NSSI provides an escape from distressing feelings and perceptions
Dependence	My NSSI will get worse My problems are so serious that NSSI is the only option	Individual	Perceptions that NSSI is the only and best option for self-regulation; that there are no alternative options; and of reliance on NSSI
Anti-dissociation	NSSI stops me feeling numb NSSI makes me feel "real" or alive	Individual	Perceptions that NSSI counters dissociation and emotional numbness
Problematic	NSSI creates a lot of problems for me I hate my NSSI injuries	Individual	Perceptions that NSSI is very counter-productive and has resulted in a range of negative consequences
Anti-suicide	NSSI is a compromise instead of killing myself I deliberately use NSSI to avoid acting on suicidal thoughts	Individual	Perceptions that NSSI is a useful and possible means of avoiding acting on suicidal thoughts
Enjoyable	NSSI is enjoyable NSSI is uplifting	Individual	Perceptions that NSSI is affectively pleasant and positive
Belonging	NSSI helps me fit in with other people NSSI helps me connect with other people	Interpersonal	Perceptions that NSSI leads to meaningful emotional connections and support from other people
Stigma	People judge and criticise my NSSI	Interpersonal	Perceptions that other people do not respond empathically to NSSI and instead judge, criticise, or punish it

Interpersonal influence	People do not understand my NSSI NSSI makes people sorry for the way they treated me NSSI makes people take my problems seriously	Interpersonal	Perceptions that NSSI is a means to rebel and prove others wrong, and a means to elicit meaningful support and care from others. It has items that appear in the SABS Eliciting help and Revenge factors

Note. NSSI = Nonsuicidal self-injury

4.8. Confirmatory Factor Analysis

Moving from EFA to Confirmatory Factor Analysis (CFA) is best practice in scale development because EFA is exploratory and does not test specific hypotheses, control for method bias, or account for the fact that some items are purer manifestations of latent variables than other items (different items have different factor loadings) (Brown, 2015). Conducting EFA and CFA on independent samples is also important because EFA may fit particular idiosyncrasies (e.g., random error) of one dataset but not generalise to other datasets (Brown, 2015). All CFAs were conducted using the R lavaan package, version 0.5-23.1097 (Rosseel, 2012).

Testing Competing Measurement Models for the SABS and NSIBS

Several potential CFA models were tested in relation to the SABS and NSIBS. These tested the generalizability of the EFA results and explored different conceptualizations of the EFA factor structures. We first tested the results of the EFAs. We specified 7 correlated factors for the SABS and 10 correlated factors for the NSIBS because these factor structures had consistently emerged in EFAs of Samples 3, 4, and 5. A correlated factors model allows the strength of correlations between factors to vary. We also examined a 9 factor model for the NSIBS whereby the items that make up the *Interpersonal influence* and *Belonging* factors were merged to form one factor. This model was specified because in Sample 4, these factors merged. Comparing a 9-factor model against a 10-factor model allowed us to test whether the *Interpersonal influence* and *Belonging* factors are best understood as two separate closely related factors or one single factor.

Three other potential measurement models were also tested for the items and factors that make up the SABS and (separately) the NSIBS. First, a bifactor model was examined in which all scale items load on a single general factor and, separately, the items that make up each subscale load on corresponding specific factors. The general and specific factors are uncorrelated and the specific factors are uncorrelated. In this model, the general factor accounts for whatever correlations are observed between factors. However, if factor correlations are unequal, the general factor is not able to equally account for the correlation between specific factors (Morgan, Hodge, Wells & Watkins, 2015). A bifactor model was specified for the SABS and separately for the NSIBS.

Next, a second-order model was examined in which specific factors (7 in the case of the SABS; 10 in the case of the NSIBS) load on a general second-order that might be thought of as representing beliefs about SA or NSSI. Last, an additional second-order model was examined which specified two general second-order factors representing self versus others.

This model was specified to account for the fact that some SABS and NSIBS subscales measure cognitions in relation to how SA and NSSI impact oneself whilst other subscales measure cognitions in relation to how SA and NSSI impact other people. Specifically, the SABS contains three factors that measure beliefs about how SA relate to oneself (*Self-punishment*, *Escape*, *Dependence; eg., "I attempt suicide because I am worthless and unlovable"*) and four factors that measure beliefs about how SA relates to other people (*Belonging, Stigma, Eliciting help, Revenge; e.g., "Attempting suicide shows other people how distressed I feel"*). The NSIBS contains seven factors that measure beliefs about how NSSI relate to oneself (*Self-punishment*, *Escape*, *Dependence*, *Anti-dissociation*, *Problematic*, *Anti-suicide, Enjoyable; eg., "NSSI helps me escape negative emotions"*) and three factors that measure beliefs about how NSSI relates to other people (*Belonging*, *Stigma*, *Interpersonal influence; e.g., "NSSI makes people take my problems seriously"*).

Testing Whether the SABS and NSIBS can be Considered Independent

We next tested whether the SABS and NSIBS can be understood as separate scales/constructs to further achieve our aim to directly contribute to the debate regarding whether it is possible or useful to conceptualise SA and NSSI as one or two constructs. Five of the subscales of the SABS and NSIBS (*Self-punishment*, *Escape*, *Dependence*, *Belonging* and *Stigma*) convey such a similar theme that these subscales were assigned the same label in each scale. These subscales contain very similar items and it is plausible that these subscales are redundant and tap the same construct. We therefore directly tested whether identically labelled subscales could be best understood as one construct or two separate correlated constructs.

We then explored whether the SABS and NSIBS scales as a whole can be best understood as measuring one or two broad constructs. One construct would mean that the items and factors that make up the SABS and NSIBS are so closely related that they need to be understood as tapping one construct. Two constructs would mean that the items and factors that make up the SABS need to be understood as tapping a separate construct to the items and factors that make up the NSIBS. We tested whether the SABS and NSIBS are best understood as one or two broad scales at the factor level, specifying 17 correlated factors (7 SABS factors + 10 NSIBS factors), and then the item level, specifying 65 items (26 SABS items + 39 NSIBS items).

Model Fit Statistics

Acceptable fit was operationalized as Comparative Fit Index (CFI) $\geq$ .90, Tucker Lewis Index (TLI) $\geq$ .90, and Root Mean Squared Error of Approximation (RMSEA) $\leq$ .08.

Good fit was operationalized as CFI $\geq$ .95, TLI $\geq$ .95, Standardized Root Mean Square Residual (SRMR) $\geq$ .08, and RMSEA $\leq$ .06 (Hu & Bentler, 1999). Our data were somewhat non-normal so we computed maximum likelihood CFA estimation with "robust" (Huber-White) standard errors and a scaled test statistic that is (asymptotically) equal to the Yuan-Bentler test statistic, which adjusts for this issue. "Robust" versions of the X^2, TLI, CFI, and RMSEA are presented, although we note that when standard versions of these fit statistics were examined, the model fit results were similar and the conclusions were identical. Competing models were compared using (i) Bayesian Information Criterion (BIC), where lower BIC statistics suggest better fit whilst adjusting for parsimony, and (ii) CFI, using a .002 cutoff (Meade, Johnson & Braddy, 2008). The BIC and CFI model comparison indices are advocated over the chi-squared and chi-squared difference statistics for judging model fit and comparing competing models because these statistics are less compromised by large sample sizes and are more sensitive to detect differences (Cheung & Rensvold, 2002; Meade et al., 2008).

Comparison of CFA Models

Several competing CFA models are presented in Table 6. The CFA results indicate that the SABS and NSIBS each consist of separate correlated subscales. A 7-factor model fit the SABS well: The CFI and TLI statistics demonstrated a fit that was between "acceptable" and "good" and the SRMR and RMSEA evidenced a "good" fit. A 10-factor model fit the NSIBS well: Again, the CFI and TLI statistics demonstrated a fit that was between "acceptable" and "good" and the SRMR and RMSEA evidenced a "good" fit. Considering the complexity of the CFA models (7 and 10 separate factors), we see these fit statistic results as impressive. The bifactor and second-order factor models were a very poor fit.

Table 6 provides clear evidence that the SABS and NSIBS are best understood as being independent from one-another (separate measures). When identically-named subscales were compared, model fit statistics consistently indicated that each subscale is best understood as being independent from the other identically-named subscale. Models in which the identically-named subscales were specified as being part of the same single construct provided a very bad fit. The models that examined whether the SABS and NSIBS can be understood as one single measure also evidenced very bad fit.

Taken together, these results indicate that the SABS consists of 7 correlated separate subscales, that the NSIBS consists of 10 correlated separate subscales, and that cognitions about SA (as operationalized by the SABS) are separate to cognitions about NSSI (as operationalized by the NSIBS). These results corroborate our earlier finding that SA and

NSSI are perceived as distinct phenomena. They also corroborate the factor structures that were observed in independent EFAs in Samples 3, 4, and 5.

Table 6 Confirmatory Factor Analytic Models of the Suicide Attempt Beliefs Scale (SABS) and Nonsuicidal Self-Injury Beliefs Scale (NSIBS)

							90% CI for RMSEA		
Model	X^2	*df*	BIC	TLI	CFI	SRMR	Lower	RMSEA	Upper
SABS									
Model 1. 7 factors	675.528***	278	62867.149	.930	.940	.053	.047	.052	.057
Model 2. Bifactor	1053.908***	273	63320.116	.864	.886	.098	.068	.072	.077
Model 3. 1 second-order factor	1104.455***	292	63302.895	.863	.877	.100	.068	.073	.077
Model 4. 2 second-order factors	988.344***	291	63165.730	.882	.895	.104	.063	.067	.072
NSIBS									
Model 1. 10 factors	1423.947***	657	91483.309	.925	.933	.055	.042	.045	.048
Model 2. 9 factors	1660.085***	666	91686.366	.904	.914	.058	.048	.051	.054
Model 3. Bifactor	1789.211***	663	91839.022	.891	.903	.082	.051	.054	.057
Model 4. 1 second-order factor	2036.916***	692	91638.414	.875	.883	.092	.055	.058	.061
Model 5. 2 second-order factors	1939.026***	691	91834.671	.884	.892	.116	.053	.056	.059
Similarly themed subscales									
Self-punishment 2 factors	116.719***	13	16791.346	.918	.949	.044	.108	.128	.150
Self-punishment 1 factor	379.438***	14	17195.270	.700	.800	.091	.224	.245	.267
Escape 2 factors	18.177***	19	20355.477	1.001	1.000	.018	.000	.000	.036
Escape 1 factor	577.477***	20	20978.399	.532	.666	.133	.205	.220	.236
Dependence 2 factors	102.847***	26	22715.648	.942	.958	.038	.057	.072	.086
Dependence 1 factor	513.206***	27	23153.547	.655	.741	.093	.161	.174	.187
Belonging 2 factors	7.733***	8	11579.490	1.00	1.00	.013	.000	.000	.052
Belonging 1 factor	275.261***	9	11892.280	.705	.823	.075	.210	.234	.258
Stigma 2 factors	54.103***	19	19922.562	.968	.978	.036	.042	.061	.080
Stigma 1 factor	472.795***	20	20387.292	.655	.753	.098	.184	.200	.216
SABS and NSIBS 1 scale									
17 factors	2021.872***	118	38055.757	.542	.602	.127	.158	.164	.171
65 items	14938.861***	2015	165339.839	.348	.369	.130	.103	.105	.106

Note. *** = $p < .001$.

4.9. General Discussion

We developed the SABS and the NSIBS across six large, independent samples of people with lived experience of SIB (total N = 3,313). Our factor analyses revealed clearly interpretable and differentiated factors that replicated well across four samples, pointing to the robustness and generalizability of our findings. Our results highlight the heterogeneous and multidimensional nature of SA and NSSI beliefs: The SABS is characterised by seven correlated factors and the NSIBS is characterised by ten correlated factors.

The development of the SABS and NSIBS present a host of intriguing findings. We briefly discuss the key findings. First, the SABS and NSIBS each contain items that measure perceptions about how SIB relates to oneself (e.g., "I attempt suicide because I am worthless and unlovable;" "NSSI helps me escape negative emotions") and items that measure perceptions about how SIB relates to other people (e.g., "Attempting suicide shows other people how distressed I feel;" "NSSI makes people take my problems seriously"). This highlights the potential importance of individual and sociocultural factors in explaining the development and repetition of SA and NSSI. The SABS and NSIBS each contain a factor we labelled *Stigma*, which taps beliefs that other people do not respond empathically to SA/NSSI and instead judge, criticise, or punish these behaviors. Endorsing this subscale, for instance, along with the *Dependence* subscale, would indicate that an individual perceives that SA and/or NSSI are their only and/or best option for self-regulation, even though these engaging in these behaviors alienates them from significant others.

Second, several subscales point to similarities between SA and NSSI. For example, we found that the SABS and NSIBS *both* contain an *Escape* factor, indicating that individuals perceive SA *and* NSSI as potentially providing temporary relief from distressing cognitive and emotional experiences. This finding is interesting because several theories posit an escape function specifically in relation to SA (e.g., Baumeister, 1990; O'Connor, 2011; Williams, 1997) or NSSI (Chapman, Gratz & Brown, 2006).

We also uncovered a *Belonging* factor in both instruments, which taps perceptions that SA and NSSI will lead to meaningful emotional connections and support from other people. The fact that some individuals (rather counterintuitively) believe that SA will potentially help them to connect with others illustrates the fact that people construct SA and NSSI idiosyncratically. This subscale is also of note given that the interpersonal theory of suicide (Joiner, 2005; Van Orden et al., 2010) predicts that *thwarted* belongingness is a key driver of SA. Some items of the SABS *Dependence* subscale (e.g., "My life would be worse without suicide attempts") especially highlight the fact that people relate to SA and NSSI

idiosyncratically. These and other somewhat counterintuitive SABS items suggest that when asked about SA, participants actually respond in relation to killing themselves and/or thinking or talking about killing oneself.

Last, the *Anti-suicide* NSIBS subscale highlights the fact that some individuals believe that NSSI is a useful and possible means of avoiding acting on suicidal thoughts. Endorsing this subscale presumably indicates a potentially protective factor against SA. However, as perceptions can quickly change, and NSSI can result in accidental death[9], we stress that *endorsing the Anti-suicide subscale cannot be interpreted as a strong or stable protective factor*.

Characterising Self-Injurious Cognitions

We argued that a major rationale for developing the SABS and NSIBS arose from the fact that there does not seem to be consensus regarding which cognitions characterise SA and NSSI. This is evidenced by the existence of fifteen or so measures of the reasons, functions, and motivations for SAs and NSSI, and the fact that each of these scales provides somewhat different content coverage. The SABS and NSIBS were developed to clarify which beliefs characterise SA and NSSI and what factor structure best explains SA and NSSI cognitions.

We generated extremely large item pools and included multiple markers of potential constructs to give meaningful content-based factors a good chance of emerging (Clark & Watson, 1995; Loevinger, 1957). It is noteworthy that we included items from existing measures of the reasons for SAs and NSSI in our item pools. Perhaps the development of the SABS and NSIBS may unify the field somewhat in its understanding of the basic constituent elements of SICs.

Item generation was guided by the broad hypothesis that SICs are characterised by positive, negative, and facilitating SICs. We revisit facilitating cognitions below. The majority of the subscales of the SABS and NSIBS appear to measure positive SICs (cognitions about the perceived advantages of SIB) and few of the subscales appear to measure negative SICs (cognitions about the perceived disadvantages of SIB). We deliberately refrained from claiming that particular subscales measure either positive or negative SICs in order to account for the complex phenomenology of SA and NSSI. It seems conceivable, for instance, that some respondents may report that a particular SABS or NSIBS item describes an advantage of SIB for them whilst other respondents describe the same item

[9] For example, lethality of SIB method can indicate suicidal intent, but not always or reliably (cf. Silverman et al., 2007a). Brown, Henriques, Sosdjan, and Beck (2004) found a minimal association between suicide intent and medical lethality

as a disadvantage of SIB for them. It also seems possible that some subscales contain positive and negative SICs. The best way to make sense of the function of particular items will probably be to ask each individual whether particular items represent an advantage or a disadvantage of SA or NSSI for them, and to test how the activation of particular items increases or decreases risk for SA or NSSI. For example, by linking activation of particular SABS and NSIBS items to changes in the strength of SA and/or NSSI urges, desire, or intent.

Conceptual Debate

Our second major goal in developing the SABS and NSIBS was to test whether SA and NSSI are best understood as one construct or two separate constructs. We presented several lines of evidence to indicate that, from a cognitions perspective, SA and NSSI are separate constructs. In Sample 1, we phrased items generically in relation to "self-injurious behavior," with no reference to suicidal desire or intent. Participants then completed every item in our item pool twice: once in relation to SA and once in relation to NSSI. We found only moderate agreement between the ratings for SA and NSSI *when rating exactly the same items*, which we took as evidence that participants perceived SA and NSSI as similar but separate constructs. We then rephrased items to refer to SA or NSSI and proceeded to develop two separate scales. We found consistent evidence across independent samples that SA cognitions evidenced a different factor structure to NSSI cognitions.

We then used CFA to directly test whether the SABS and NSIBS can be understood as separate scales. These analyses were particularly important in light of the fact that five of the subscales of the SABS and NSIBS (*Self-punishment*, *Escape*, *Dependence*, *Belonging* and *Stigma*) convey such a similar theme that they were assigned an identical label in each scale. Taken together, these analyses revealed clear evidence that the SABS and NSIBS are best understood as related but separate measures. The implications of these findings are far-reaching as they support the use of separate terminology and definitions of SA and NSSI, and indicate that SA and NSSI need to be separated in research designs. Although SA and NSSI often co-occur in the same individual (Klonsky, May & Glenn, 2013; Nock et al., 2006) and relate to one-another (see Table 1), SA and NSSI are separate. Whether or not SA and NSSI are separate is analogous to the use of different illicit drugs. For example, a regular drug user could quite conceivably hold the same belief about two different drugs, believing that cannabis and heroin both help them form meaningful social connections. However, as a result of whatever internal or external circumstances, they may strongly endorse this particular belief for heroin and only mildly endorse it for cannabis (perhaps their current social circle are heavy heroin users). The person may also use heroin more than cannabis. Just like beliefs

about cannabis versus heroin, our analyses demonstrate that SA and NSSI cognitions are separate – a fact which does not mean SA and NSSI cannot co-occur or relate to one-another (the evidence demonstrates they often do) or be underpinned by the same mechanisms.

Clinical Utility

We are hopeful that the SABS and NSIBS will be useful to researchers and clinicians alike. When used clinically, we advise that the SABS and NSIBS are used to facilitate collaborative, close questioning regarding what SA and/or NSSI idiosyncratically *means* to each individual. This information can be coupled with theory and other clinical and contextual information to make individualised predictions about risk and to formulate targeted therapeutic interventions. Clinicians are advised to explore *why* respondents endorse particular items of the SABS and NSIBS and to link SA and NSSI cognitions to other presenting problems and contextual information. The SABS may help individuals clarify whether they want to die or whether they actually want their suffering to end.

We have included a section at the end of the SABS and NSIBS for respondents to record important thoughts or beliefs about SA and NSSI that are not included in each scale. Given the extensive nature of our item generation and the size of our item pools, it is likely that any thoughts or beliefs that are reported will be similar to items that featured in our item pools. If this is the case, the clinician will be able to glean an idea of which factor particular cognitions are likely to stem from.

Limitations and Future Directions

We see our results as very promising but further research is of course needed to explicate the construct validity of these instruments more fully. The factor structures we reported here seem fairly robust and generalizable. Nevertheless, they must be replicated in other samples before they can be considered a definitive account of the structure of our scales (e.g., clinical samples, ethnically-diverse samples, samples with different comorbidities). Future research particularly needs to examine the measurement invariance of the scales and validate their use in different populations. We retained a subset of items from each factor (see Tables 3 and 4) because this seemed most appropriate given the mental states of people who are considering SA and NSSI. However, researchers who are interested in focusing on particular subscales are advised to use the full set of items from each subscale as these will of course be more internally consistent and tap each construct more completely.

It will be useful to develop child and adolescent versions of the SABS and NSIBS as well as brief versions of each instrument that can be used in epidemiological research or clinical settings where a screening measure is more desirable and practical. We are in the

process of developing these instruments. The development of a child and adolescent version of the SABS and NSIBS will help clarify whether the same cognitions characterise SA and NSSI cognitions across the lifespan and may shed further light on the onset and development of SA and NSSI, perhaps with implications for prevention and early intervention. Since a thorough assessment of SIB would ideally involve informant reports, it would be interesting for informant versions of the SABS and NSIBS to be developed, and for research to explore the explanatory value of self-informant discrepancies (Connelly & Ones, 2010; De Los Reyes & Kazdin, 2005; Fitzsimons, Finkel & vanDellen, 2015).

Another important area for future research involves exploring the potential existence and role of facilitating cognitions. Facilitating cognitions failed to emerge in our structural analysis as a distinct factor. However, these items were frequently endorsed for both SA and NSSI (see Supplementary material for Chapter 4), potentially indicating that they have a role to play in both sets of behaviors. We believe that facilitating cognitions did not emerge in our structural analyses because they likely take the form of "automatic thoughts" (Beck et al., 1979) and will therefore be more idiosyncratic and situation and person-specific. In contrast, the SABS and NSIBS appear to measure relatively enduring (see test retest reliability in Part II), widely endorsed personal meanings for SIB that confer vulnerability across SIB individual differences and samples. The fact that facilitating cognitions did not emerge in our structural analyses is also consistent with the notion that they originate from positive beliefs about SA or NSSI (cf. Beck, Wright, Newman & Liese, 1993). This seems a plausible hypothesis because people do not give themselves permission to engage in behaviors – especially self-destructive behaviors – which have no personal meaning and value.

Finally, we recognise that our data are exclusively self-report and therefore vulnerable to the limitations associated with using a single method (e.g., inflated estimates of associations between constructs; Podsakoff, MacKenzie, Lee, & Podsakoff, 2003). It will be useful for future research to validate the SABS and NSIBS against data collected using other methods, such as informant and clinician ratings.

4.10. Supplementary Material for Chapter 4

Item Generation and Refinement

Strenuous efforts were made to generate an exhaustive list of distinct cognitions about SIB, and to incorporate the perspectives of clinicians, researchers, and individuals with lived experience of SIB during the different stages of item generation and refinement (see Figure 1 in the main article). This methodology aimed to make the scales that were developed clinically useful, understandable, and broadly relevant; and to minimise the potential influence of researcher bias during the scale development process. Item generation and refinement involved the following steps.

Thought Capture Item Generation

A "thought capture" qualitative study was conducted in order to elicit a broad and representative pool of self-injurious cognitions (SICs) from individuals with lived experience of SIB. Attempts were made to elicit a wide range of types of SICs, including thoughts, assumptions, beliefs, expectations, reasons, attitudes, imagery, dreams, nightmares, memories, and meta-cognitions.

The questions used to elicit SICs in the thought capture exercise were based on a range of methodologies that have been used previously in clinical and health psychology to elicit cognitions (e.g., Ajzen & Driver, 1991, 1992; Beck, 1995; Beck, Rush, Shaw & Emery, 1979; Butler, Fennell & Hackmann, 2010; Miller & Rollnick, 2002; Pearson et al., 2013; Wells, 1997, 2000; Winter, 2003). The questions were independently reviewed by five National Health Service (NHS) Clinical Psychologists and ten members of the public to verify that they were clear and that the protocol was likely to elicit the types of information that the research team expected. The questions were further refined based on this feedback.

Participants were instructed that they would be asked a series of questions about their self-injurious thoughts. To avoid fatigue and frustration, participants were explicitly instructed that they could skip questions as needed in order to avoid repeating themselves and because some of the questions may not be relevant to every person. As we were interested in exploring which cognitions characterise any and all types of SIB, and so as not to artificially constrain responses, we did not ask participants to differentiate by method of SIB or whether the SIB was motivated by suicidal thoughts or intent.

To avoid confusion and in the interests of eliciting specific and nuanced SICs, participants were provided with a definition of self-injurious behaviour (SIB) and a clarification of behaviours which we do not consider to be SIB. SIB was defined as

"physically hurting or injuring yourself on purpose, whether you intend to kill yourself or not. (Also called suicide, self-injury, self-harm)."

63 thought capture questions were administered. The questions were initially broad and open-ended in the aim of minimising the influence of the research team's position on the information reported. Participants were asked to report any thoughts, feelings, mental images, dreams or nightmares, memories, and other experiences "(e.g., voices, sounds, smells, sensations, tastes, etc)" related to SIB and its sequalae, including cognitions about physical pain, blood, and injuries experienced as part of SIB. Thereafter, the questions became progressively more specific and guiding in order to comprehensively elicit cognitions in relation to each of the three hypothesised domains of SICs (positive SICs, negative SICs, and facilitating SICs). The survey ended with a "mop-up" question which asked participants to record anything about self-injurious thoughts or behaviour that they had not reported so far.

Responses were used to form items that were rationally organised into homogeneous item composites (HICs; Hogan, 1983) to ensure comprehensive coverage of different cognitive themes. This procedure is recommended because having multiple markers (at least 5 items) for each cognitive theme ensures that a corresponding factor has a reasonable chance to emerge during structural analyses. This procedure also helps avoid the inclusion of duplicate items in item pools.

Participants

Individuals who have thought about or engaged in any form of SIB in the last six months were recruited online. They provided responses to the thought capture exercise using the survey software Qualtrics. Participants were entered into a prize draw for a chance to win a £100 Amazon voucher.

Research Team Item Generation

Independent of the thought capture study, the research team generated a broad item pool explicitly designed to operationalize positive, negative, and facilitating SICs. Potential markers of the three domains of SICs were generated by drawing on a broad range of sources: (1) existing measures of the reasons, functions, and motivations for engaging in or refraining from SA and/or NSSI (Beck & Steer, 1991; Holden, Kerr, Mendonca & Velamoor, 1998; Kleindienst et al., 2008; Klonsky & Glenn, 2009; Lewis & Santor, 2008; Linehan, Goodstein, Nielsen & Chiles, 1983; Lloyd et al.,1997; Ma & Klonsky, 2013; Matson et al., 1999; Nock et al., 2007; Osman et al., 1998; Osuch et al., 1999; Santa Mina et al., 2006; Turner, Chapman & Gratz, 2014; Whitlock et al., 2014); (2) existing measures of perceptions that are thought to underlie suicidal thoughts and/or SA specifically, such as hopelessness,

defeat/entrapment, burdensomeness, unlovability, unbearability, unsolvability, thwarted belongingness, and acquired capability (Beck et al. 1974; Rudd et al., 2015; Van Orden et al., 2008); (3) themes that have emerged from previous qualitative research on SIB (e.g., Adams, Rodham & Gavin, 2005; Brown & Kimball, 2011; Harris, 2000); (4) indications regarding SICs in the SIB literature; (5) online social support forums and YouTube testimonials; (6) the clinical experience of the first author; (7) cognitive-behavioural theory, research, and measures for other self-destructive behaviours, especially binge eating and substance misuse; (8) clinical, health, and social psychology theory and research on behaviour, health behaviour, behaviour change, and risk perception; and (9) research regarding transdiagnostic cognitive processes and metacognition (e.g., rumination, thought suppression, intrusive thoughts and memories).

All items were phrased to describe any form of cognition *about* SIB, irrespective of suicidal intent (i.e. items did not differentiate bwteeen SA and NSSI). No reference was made to the method of SIB. Items were rationally organised into HICs.

Procedure for Combining Item Pools

A revised item pool was created by combining the item pool that emerged from the thought capture exercise with the item pool that had been generated by the research team. As with the previous item pools, each item was phrased in relation to "SIB" (self-injurious behaviour), rather than specifying suicidal intent or particular SIB methods. When an item from each item pool expressed the same content and meaning, the clearest, simplest, or shortest item was retained for the revised item pool. All items that expressed unique content were included in the new, combined item pool and the research team put no boundaries, conditions, or hypotheses on which items should or should not be retained. Although items were eventually discarded based on psychometric analyses, beginning the process with an over-inclusive item pool ensured that the final instrument/s adequately cover all content areas (Clark & Watson, 1995; Loevinger, 1957).

Item Refinement

The next step in the scale development was for the combined item pool to be (simultaneously) reviewed and refined by three SIB stakeholder groups: (1) Public and private sector clinicians from across the world who worked in specialist SIB services or services that regularly encountered SIB (e.g., personality disorder treatment services), (2) researchers and clinician-academics who have recently (last 5 years) published on SIB, self-destructive behaviour, or problematic health behaviours, and (3) People with lived experience of SIB who have "thought about self-harm, nonsuicidal self-injury, or suicide in the last two

years." Researchers and clinicians were recruited through speculative emails. People with lived experience of SIB were recruited online through the same methods as all participants were recruited (see main article).

Participants were not paid. As with the item pool generated by the research team, this facet of the scale development process was guided by our underlying hypothesis that the presence and interaction of positive SICs, negative SICs, and facilitating SICs characterise SA and NSSI. Items were therefore grouped according to the facet of the hypothesis they appeared to resemble. Participants were then provided with operational definitions of the three SIC domains and asked to rate (Yes/No) whether each item matched the operational definition. The following definitions were provided:

Positive beliefs about SIB are beliefs about the individual and interpersonal advantages of SIB. These beliefs are what makes SIB sometimes seem an attractive, favourable, or valued option. These beliefs make SIB more likely to occur. (The word "positive" does not imply that SIB is seen as a "helpful" or "constructive" option; only that it is sometimes seen as an advantageous or desirable option. SIB may in fact be a person's least bad or only option). Do the following items fit this definition of INDIVIDUAL "positive beliefs about SIB" / Do the following items fit this definition of INTERPERSONAL "positive beliefs about SIB"?

Negative beliefs about SIB are beliefs about the individual and interpersonal disadvantages and detrimental consequences of SIB. These beliefs describe how SIB can be perceived as an aversive, unfavourable, or undesirable option. These beliefs make SIB less likely to occur. Do the following items fit this definition of INDIVIDUAL "negative beliefs about SIB?"" / INTERPERSONAL "negative beliefs about SIB?""

"Facilitating thoughts justify and give "permission" to engage in SIB. Facilitating thoughts occur in response to conflicting positive and negative beliefs about SIB and make SIB more likely to occur. They activate and strengthen positive beliefs about SIB and inhibit and weaken negative beliefs about SIB. Facilitating thoughts enable individuals to proceed with a behaviour that they know to be unconstructive. Do the following items fit this definition of "facilitating thoughts about SIB?"

Because the SIB literature is like no other in the extent to which different people interpret different terminology to mean different things, definitions of what SIB is and is not were provided on every page of each online survey. The following definitions and clarifications were provided to each Sample:

Self-injurious behaviour = physically hurting or injuring yourself on purpose. It includes suicide attempts and nonsuicidal self-injury.

A suicide attempt = intentionally physically injuring yourself in order to kill yourself.

Nonsuicidal self-injury = intentionally physically injuring yourself, but with no desire or intention of killing yourself or being dead.

Things that are <u>not</u> a suicide attempt or nonsuicidal self-injury:

Behaviours which unintentionally cause physical harm long-term (e.g. smoking, over-eating, binge drinking, eating disorders, unprotected sex)

Accidentally injuring yourself (e.g. accidentally touching something hot)

Behaviours which change your body for a cultural reason (e.g. body piercing, tattooing)

A definition of one other domain of SICs termed "anti-suicide cognitions" was also provided to Sample 1. These items were hypothesised to be a facet of individual positive beliefs about SIB and described using NSSI as a means to avoid acting on suicidal thoughts and urges. In contrast to the rest of the item pool (which was phrased in reference to SIB), anti-suicide items specified NSSI and SA as separate phenomenon and used the terminology of "NSSI" and "suicide". The following clarification was provided:

This section asks about a group of thoughts that are specific to nonsuicidal self-injury (NSSI). To recap, NSSI involves physically hurting or injuring yourself on purpose with no intention or expectation at all of dying. Anti-suicide NSSI thoughts are thoughts about engaging in nonsuicidal self-injury as a way to avoid acting on suicidal thoughts or urges (i.e. to avoid making a suicide attempt). Do the following items fit this definition of "anti-suicide NSSI thoughts"?

After this exercise, participants in the item refinement stage (researchers, clinicians, and people with lived experience of SIB) were asked to indicate whether they believed that any self-injurious thoughts were missing from the item pool, whether any items could be rephrased to be simpler or clearer, and whether any items should be deleted. Participants were also given the opportunity to make comments. Several changes were made to the item pool in response to the feedback received.

Development of the Final Scales: Samples 3, 4, and 5

Sample 3

In Sample 3, 92 SA items and 131 NSSI items were administered.

Results

11 highly correlated items were deleted from the NSSI item pool. PAs indicated 11 SA factors and 12 NSSI factors.

SA Item Pool. An EFA suggested the existence of 7 interpretable factors, which we labelled *Dependence* (25 items; e.g., "Attempting suicide is one of the best things in my life"), *Stigma* (13 items; e.g., "People judge and criticise my suicide attempt(s)"), *Interpersonal influence* (7 items; e.g., "Attempting suicide makes people take my problems seriously"), *Escape* (11 items; e.g., "Attempting suicide clears my mind"), *Belonging* (10 items; e.g., "Attempting suicide helps me fit in with other people"), *Self-punishment* (8 items; e.g., "I attempt suicide to show how much I hate myself"), and *Protest* (8 items; e.g., "Attempting suicide is a way to get back at people who have hurt me"). All of these factors were apparent in Sample 1.

NSSI Item Pool. An EFA suggested the existence of 10 interpretable, clear factors, which we labelled *Escape* (27 items; e.g., "NSSI helps me escape negative emotions"), *Dependence* (15 items; e.g., "I rely on SIB"), *Interpersonal influence* (13 items; e.g., "NSSI helps me get revenge against others"), *Stigma* (7 items; e.g., "People judge and criticise my SIB"), *Problematic* (12 items; e.g., "NSSI makes my problems worse"), *Enjoyable* (e.g., "NSSI is enjoyable"), *Anti-suicide* (6 items; e.g., "NSSI is a replacement for suicidal behaviour"), *Self-punishment* (6 items; e.g., "I engage in SIB to punish myself"), *Belonging* (8 items; e.g., "NSSI helps me fit in with other people"), and *Anti-dissociation* (6 items; e.g., "NSSI takes me out of a detached state").

Distress and Impaired Functioning Items. The diagnostic criteria for most psychiatric diagnoses includes the presence of functional impairment and/or clinically significant distress. Several self-report measures that aim to have relevance for the Diagnostic and Statistical Manual of Mental Disorders include items that measure functional impairment and clinically significant distress (e.g., the Posttraumatic Diagnostic Scale for DSM–5; Foa et al., 2015). We explored the effect of adding distress and impaired functioning items to each of the SA and NSSI item pools, copying the phrasing used by Foa et al. (2015). The distress item in each item pool asked "How much have these beliefs about SA/NSSI been bothering you?" The impaired functioning item in each item pool asked "How much have these beliefs about SA/NSSI been interfering with your everyday life (for example relationships, work, or

other important activities)?" Both items were rated on the same 7-point Likert scale (ranging from *Strongly agree* to *Strongly disagree*) as that used for the rest of the item pool. These items were added to Samples 2 and 3.

In Sample 2, we found that the two items tapping distress and impaired functioning performed poorly. They demonstrated negative factor loadings for SA and NSSI. They formed their own factor for SA and joined two items for NSSI to make a factor with no obvious theme. The two distress and impaired functioning items were administered to Sample 3 to further explore the properties of these items for SA and NSSI. These items again had negative factor loadings. They formed their own factor for SA and loaded on the *Dependence* factor for NSSI. The poor performance of these items across Samples 2 and 3 meant that they were omitted from Samples 4 and 5 and do not feature in the final versions of the SABS or NSIBS.

Discussion

Sample 3 provided encouraging results: a large number of interpretable factors emerged clearly for SA and NSSI. These results were important given the question of which factor structure to retain from Samples 1 and 2. The results observed in Sample 3 provided further confidence (on top of Sample 2) that the structural results observed in Sample 1 were probably not confounded by the design or by the administration of a large item pool. Nevertheless, we recognised that the task of developing scales that reliably measure a large number of distinct factors is psychometrically challenging and collected an additional independent sample to explore whether the factor structure observed in Sample 3 would replicate.

Sample 4

At this point, we shifted our emphasis from sensitivity (administering overly large item pools to encourage distinct potential factors to emerge) to specificity (only including strong markers of a lesser number of clear factors). In Sample 4, we administered items that demonstrated reasonably strong loadings (<.40) and that did not cross-load, as well as four items that had inadvertently been omitted from Sample 3 ("NSSI is a compromise instead of killing myself," "NSSI is a way to avoid suicidal urges," "I engage in NSSI because I deserve to suffer," "NSSI stops me feeling numb"). In Sample 4, 40 SA items and 61 NSSI items were administered.

Results

3 highly correlated items were deleted from the SA item pool and 1 highly correlated item was deleted from the NSSI item pool. PAs indicated 7 SA factors and 9 NSSI factors.

SA Item Pool. An EFA suggested the existence of 7 interpretable, clear factors, which we labelled: *Belonging* (4 items; e.g., "Attempting suicide makes me feel part of a group"), *Stigma* (6 items; e.g., "People judge and criticise my suicide attempt(s)"), *Escape* (6 items; e.g., "Attempting suicide stops upsetting thoughts going round and round in my mind"), *Self-punishment* (5 items; e.g., "I attempt suicide because I deserve to suffer"), *Interpersonal influence* (5 items; e.g., "Attempting suicide makes people take my problems seriously"), *Dependence* (6 items; e.g., "Attempting suicide is the only method of coping that works for me"), and *Protest* (5 items; e.g., "Attempting suicide is a form of rebellion"). In contrast to previous EFAs, the PA and EFA results concurred for Sample 4. The same factor structure emerged across Samples 3 and 4.

NSSI Item Pool. An EFA suggested the existence of 9 interpretable, clear factors. The first factor involved the *Interpersonal influence* factor (e.g., "My NSSI persuades other people to change their mind") merging with the *Belonging* (e.g., "NSSI helps me fit in with other people") factor, presumably because the items were so closely themed. The remaining factors were: *Anti-dissociation* (7 items; e.g., "NSSI stops me feeling detached from myself"), *Self-punishment* (7 items; e.g., "I engage in SIB to punish myself"), *Dependence* (6 items; e.g., "I cannot cope without NSSI"), *Anti-suicide* (5 items; e.g., "NSSI is a replacement for suicidal behaviour"), *Problematic* (6 items; e.g., "NSSI makes my problems worse"), *Stigma* (6 items; e.g., "People judge and criticise my SIB"), *Enjoyable* (7 items; e.g., "NSSI is enjoyable"), and *Escape* (4 items; e.g., "NSSI helps me escape negative emotions").

As with the SA items, in contrast to previous EFAs, the PA and EFA results concurred for Sample 4. The same factor structure emerged across Samples 3 and 4, except, as noted, in Sample 4 the *Interpersonal influence* and *Belonging* factors combined to form one factor.

Sample 5

The PA and EFA results observed in Samples 3 and 4 provide evidence that a robust, replicable, and clearly interpretable factor structure has emerged to characterise SA and NSSI cognitions. Sample 5 analyses were conducted as an additional check of the generalizability of the structural results that had been observed in Samples 3 and 4. In Sample 5, 41 SA items and 65 NSSI items were administered.

Results

2 highly correlated items were deleted from each of the SA and the NSSI item pools. PAs indicated 7 SA factors and 10 NSSI factors.

SA Item Pool. An EFA suggested the existence of 7 interpretable, clear factors, which we again labelled: *Belonging* (5 items; e.g., "Attempting suicide helps me fit in with other people"), *Stigma* (6 items; e.g., "People think that my suicide attempt(s) are selfish"), *Self-punishment* (6 items; e.g., "I attempt suicide because I deserve to suffer"), *Interpersonal influence* (5 items; e.g., "Attempting suicide makes people take my problems seriously"), *Escape* (6 items; e.g., "Attempting suicide changes the way that I am thinking"), *Dependence* (6 items; e.g., "Attempting suicide is the only option I have for solving my problems"), and *Protest* (5 items; e.g., "Attempting suicide is a way to get back at people who have hurt me"). As we found in Sample 4, the PA and EFA results concurred for Sample 5. The same factor structure emerged in Sample 5 as had been observed in Samples 3 and 4.

NSSI Item Pool. An EFA suggested the existence of 10 interpretable, clear factors, which we again labelled: *Escape* (8 items; e.g., "NSSI helps me escape negative emotions"), *Self-punishment* (7 items; e.g., "I engage in NSSI because I deserve to suffer"), *Anti-dissociation* (6 items; e.g., "NSSI stops me feeling numb"), *Interpersonal influence* factor (6 items; e.g., "NSSI is a way to intentionally upset other people"), *Stigma* (6 items; e.g., "People reject me because of my NSSI"), *Dependence* (6 items; e.g., "I cannot cope without NSSI"), *Problematic* (6 items; e.g., "NSSI makes my problems worse"), *Anti-suicide* (4 items; e.g., "NSSI is a compromise instead of killing myself"), *Enjoyable* (6 items; e.g., "NSSI is enjoyable"), and *Belonging* (5 items; e.g., "NSSI helps me fit in with other people"). The *Interpersonal influence* and *Belonging* factors did not merge in Sample 5. As we found in Sample 4, the PA and EFA results concurred for Sample 5. The same factor structure emerged in Sample 5 as had been observed in Samples 3 and 4, with the only caveat being that in Sample 4 the *Interpersonal influence* and *Belonging* factors merged to form one factor.

Table S1. Descriptive Statistics for Facilitating Cognition Items Across Three Samples

	Study 1				Study 2				Study 3			
	Suicide Attempt		NSSI		Suicide Attempt		NSSI		Suicide Attempt		NSSI	
	M	SD	*M*	SD	*M*	SD	*M*	SD	*M*	SD	*M*	SD
I give myself permission to XXX	4.27	2.21	3.29	2.08	4.19	2.17	3.32	1.98	4.07	2.22	3.28	2.03
I vow that 'This will be the last time I XXX	4.22	2.40	4.43	2.32	-	-	-	-	-	-	-	-
I think things that sabotage my efforts to avoid XXX	4.09	2.02	3.78	2.01	-	-	-	-	-	-	3.74	1.92
I think things that make XXX more likely to happen	3.82	2.10	3.51	2.09	-	-	-	-	-	-	-	-
I have good reasons for XXX	3.33	2.19	2.94	1.99	-	-	-	-	-	-	-	-
I encourage myself to XXX	5.41	1.94	5.10	2.02	-	-	-	-	-	-	-	-
I tell myself that I can XXX if I do something to make up for it afterwards	6.00	1.56	5.67	1.79	5.99	1.49	-	-	6.01	1.49	5.37	1.93
It's acceptable to XXX if I do it in a particular way	5.00	2.04	4.17	2.15	5.35	1.79	4.34	2.09	5.17	1.95	4.34	2.13
SIB is more acceptable if I have opportunity to do it	5.01	1.95	4.61	2.01	5.09	1.87	4.44	1.94	4.80	1.96	4.38	2.01
I try not to think about the disadvantages of XXX	3.99	2.25	3.48	2.11	-	-	-	-	3.93	2.19	3.33	2.03
I ignore problems associated with XXX	4.01	2.12	3.56	2.04	-	-	-	-	-	-	3.35	1.90
I under-estimate the consequences of XXX	3.88	2.24	3.56	2.08	-	-	-	-	-	-	-	-
I try to ignore the physical pain that comes with XXX	3.99	2.27	4.10	2.34	-	-	-	-	-	-	-	-
I tell myself that I deserve the benefits of XXX	4.52	2.21	4.29	2.22	-	-	4.13	2.15	-	-	4.06	2.17
I find ways to justify XXX to myself	3.71	2.24	3.22	2.13	-	-	3.17	2.04	-	-	3.11	2.04
It's acceptable to XXX if I'm really upset	4.45	2.17	3.49	2.11	-	-	3.68	2.16	-	-	3.53	2.09
The benefits of XXX are worth the risks	4.48	2.26	3.53	2.09	-	-	3.67	2.04	-	-	3.60	2.07
I think that XXX won't be dangerous as long as I am careful	5.11	2.14	3.10	2.10	5.57	1.89	-	-	5.55	1.85	2.84	1.99
XXX is a problem for some people but it won't be for me	5.02	1.98	4.75	2.08	-	-	-	-	4.44	1.98	-	-
XXX is more acceptable if I do something to make up for it afterwards	5.82	1.61	5.61	1.73	5.81	1.57	-	-	5.77	1.63	5.20	1.92
I allow myself to delay when I'll XXX, knowing that I will do it later	4.15	2.12	3.74	2.10	-	-	-	-	-	-	-	-

Note. XXX denotes the use of different terminology in different Samples: In Sample 1, "SIB" was used; in Samples 2 and 3, "attempt/ing suicide" and "NSSI" were used.

CHAPTER 5

5. Characterizing Self-Injurious Cognitions Part II: Validation of the Suicide Attempt Beliefs Scale (SABS) and the Nonsuicidal Self-Injury Beliefs Scale (NSIBS).

5.1. Abstract

This article reports the validation of the Suicide Attempt Beliefs Scale (SABS) and Nonsuicidal Self-Injury Beliefs Scale (NSIBS), two new measures of beliefs about suicide attempts (SA) and nonsuicidal self-injury (NSSI). Both instruments demonstrated moderate to excellent test retest reliability over 2-4 weeks and strong internal consistency. The subscales of the SABS and NSIBS each measure relatively specific domains of self-injurious cognitions (SICs) and each instrument as a whole taps a range of distinct, moderately related content. Many of the SABS subscales demonstrate almost no relationship with the NSIBS subscales. A range of correlations support the convergent and divergent validity of the SABS and NSIBS: both instruments demonstrated small to moderate correlations with a range of clinical variables, measures of well-being, and purportedly similar, existing constructs, including the reasons for SA and NSSI, reasons to live, current suicidal thinking, and perceptions of unlovability, unbearability, burdensomeness, and thwarted belongingness. The SABS significantly incrementally predicts previous and current SICs and behavior over and above demographic characteristics, clinical, and well-being variables. People increasingly endorsed the SABS and NSIBS the more they had progressed from thoughts to taking action to engaging in multiple episodes of SA or NSSI, and the more recently they had made a SA or engaged in NSSI. Taken together, these analyses supplement those presented in our Development article and further demonstrate that the SABS and NSIBS have promising psychometric properties and need to be understood as related but separate measures: that beliefs about SA and NSSI are similar but distinct.

Chapter 5 will be submitted for peer-review to *Psychological Assessment* as: Siddaway, A. P., Wood, A. M., O'Carroll, R. E., & O'Connor, R. C. Characterizing Self-Injurious Cognitions Part II: Validation of the Suicide Attempt Beliefs Scale (SABS) and the Nonsuicidal Self-Injury Beliefs Scale (NSIBS).

5.2.Introduction

This article (Part II) is the second of two companion articles that describe the development and validation of the Suicide Attempt Beliefs Scale (SABS) and the Nonsuicidal Self-Injury Beliefs Scale (NSIBS). Part I described the development of these scales. This article explicates the internal consistency, test retest reliability, convergent, discriminant, criterion, and incremental validity of the SABS and NSIBS.

The SABS was developed to measure beliefs about SA. It consists of a 7-factor structure, with three factors measuring perceptions about how SA relates to oneself (e.g., "I attempt suicide because I am worthless and unlovable") and 4 factors measuring perceptions about how SA relates to other people (e.g., "Attempting suicide shows other people how distressed I feel"). The 7 factors of the SABS are called: *Belonging* (3 items; e.g., "Attempting suicide helps me fit in with other people"), *Stigma* (4 items; e.g., "People think that my suicide attempt(s) are selfish"), *Self-punishment* (3 items; e.g., "I attempt suicide because I deserve to suffer"), *Eliciting help* (3 items; e.g., "Attempting suicide makes people take my problems seriously"), *Escape* (4 items; e.g., "Attempting suicide changes the way that I am thinking"), *Dependence* (5 items; e.g., "Attempting suicide is the only option I have for solving my problems"), and *Revenge* (4 items; e.g., "Attempting suicide is a way to get back at people who have hurt me").

The NSIBS was developed to measure beliefs about engaging in NSSI. It consists of a 10-factor structure, with seven factors measuring perceptions about how engaging in NSSI relates to oneself (e.g., "NSSI helps me escape negative emotions") and three factors measuring perceptions about how NSSI relates to other people (e.g., "NSSI makes people take my problems seriously"). The 10 factors of the NSIBS are called: *Escape* (4 items; e.g., "NSSI helps me escape negative emotions"), *Self-punishment* (4 items; e.g., "I engage in NSSI because I deserve to suffer"), *Anti-dissociation* (3 items; e.g., "NSSI stops me feeling numb"), *Interpersonal influence* (5 items; e.g., "NSSI is a way to intentionally upset other people"), *Stigma* (4 items; e.g., "People reject me because of my NSSI"), Dependence (4 items; e.g., "I cannot cope without NSSI"), *Problematic* (5 items; e.g., "NSSI makes my problems worse"), *Anti-suicide* (3 items; e.g., "NSSI is a compromise instead of killing myself"), *Enjoyable* (4 items; e.g., "NSSI is enjoyable"), and *Belonging* (3 items; e.g., "NSSI helps me fit in with other people").

The literature regarding self-injurious behavior (SIB) is afflicted by a debate regarding whether intentional self-injury in order to die (which we refer to as a "suicide attempt; SA") is different from intentional self-injury with no intent to die (which we refer to

as “nonsuicidal self-injury; NSSI) (see Muehlenkamp, 2014; Posner, Brodsky, Yershova, Buchanan & Mann, 2014; Silverman et al., 2007a, b). A central point of contention concerns whether SA and NSSI are one or two constructs (Kapur, Cooper, O’Connor & Hawton, 2013). In Part I, we attempted to move this debate forwards by directly testing whether SA and NSSI are perceived as distinct phenomena and by exploring whether different cognitions characterise SA and NSSI. We presented several lines of evidence to indicate that SA and NSSI are perceived as similar but ultimately separate phenomena, which justified the development of two separate scales. Intriguingly, we found that SA and NSSI are characterised by some very similar beliefs and some beliefs that are specific to either SA or NSSI.

This article describes the psychometric properties of the SABS and NSIBS and presents further evidence regarding the similarities and differences between SA and NSSI beliefs. We examine the test retest reliabilities of the SABS and NSIBS over 2-4 weeks to see whether these instruments might be understood as measures of beliefs (relatively enduring personal meanings for SA and NSSI that confer vulnerability across situations but which also potentially fluctuate or change due to a range of factors; Beck, Rush, Shaw & Emery, 1979). Our examination of the internal consistency of the SABS and NSIBS subscales and total scores will be instructive in clarifying whether these instruments measure relatively narrow or relatively broad facets of SA and NSSI beliefs.

It will be informative to see how strongly the subscales of the SABS and NSIBS correlate; that is, how closely related cognitions about SA are to cognitions about NSSI. These analyses are especially important given that five subscales (*Self-punishment, Escape, Dependence, Belonging and Stigma*) convey such a similar theme in the SABS and NSIBS that they were assigned the same label. Anything other than strong correlations between the SABS and NSIBS would provide further evidence that SA and NSSI beliefs need to be considered separate, justifying the need for separate measures of SA and NSSI beliefs.

We were very keen to provide stringent tests of the need for the SABS and NSIBS. We therefore examine the convergent, discriminant, criterion, and incremental validity of the SABS and NSIBS against (i) purportedly similar, existing constructs (the reasons for SA and NSSI, reasons to live, current suicidal thinking, and perceptions of unlovability, unbearability, burdensomeness, and thwarted belongingness); (ii) related clinical constructs (depressive symptoms, anxiety symptoms, symptoms of BPD, experiential avoidance, perceived stress, difficulties in regulating positive and negative emotions, emotional reactivity); and (iii) several measures of well-being (optimism, perceived social support,

satisfaction with life, and subjective happiness and vitality), which would be expected to share less variance with other predictors in statistical models.

5.3. Method

5.3.1. Participants and Procedure

The SABS and NSIBS were developed across six independent samples that were collected online. Detailed information regarding recruitment and participant characteristics is presented in the companion article (Part I). Participants in four samples (Samples 3-6[10]) completed the SABS, NSIBS, and a range of other measures, as detailed next.

5.3.2. Measures

Self-Injurious Thoughts and Behavior. All participants were asked questions about the presence, lifetime frequency, and recency of NSSI thoughts and behavior, suicidal thoughts and behavior, and whether they had experienced thoughts about SA and NSSI at the same time or in relation to one-another (see Part I).

Suicide Attempt Beliefs Scale (SABS; Siddaway et al., in submission). The SABS measures beliefs about attempting suicide. It consists of 26 items and 7 factors (described above). Participants are asked to report how much they currently agree with each belief about attempting suicide using a 7-point Likert scale ranging from *Strongly disagree* to *Strongly agree*. In an attempt to minimize misinterpretation, a definition of "attempting suicide" is provided ("intentionally physically injuring yourself because you want and expect to kill yourself") and several behaviors that are not considered a SA, including NSSI, are defined. There is space to record "any other thoughts or beliefs you have about attempting suicide" after the NSIBS items.

Nonsuicidal Self-Injury Beliefs Scale (NSIBS; Siddaway et al., in submission). The NSIBS measures beliefs about engaging in NSSI. It consists of 39 items and 10 factors (described above). Participants are asked to report how much they currently agree with each belief about NSSI using a 7-point Likert scale ranging from *Strongly disagree* to *Strongly agree*. In an attempt to minimize misinterpretation, a definition of "NSSI" is provided ("intentionally physically injuring yourself, but with no desire or intention of killing yourself or being dead") and several behaviors that are not considered NSSI, including SAs, are defined. There is space to record "any other thoughts or beliefs you have about NSSI" after the NSIBS items.

5.3.3. Measures Completed by Sample 3

[10] Samples 1 and 2 were purely used to develop the SABS and NSIBS

Inventory of Depression and Anxiety Symptoms (IDAS) – Dysphoria subscale (IDAS-Dysphoria; Watson et al., 2007). This 10 item subscale assesses the nonspecific emotional and cognitive symptoms shared by depression and various anxiety problems. It contains single items assessing depressed mood, anhedonia, worry, worthlessness, guilt, and hopelessness, as well as two items apiece tapping psychomotor disturbance (one reflecting retardation, the other agitation) and cognitive problems. Participants indicate the extent to which they have experienced each symptom "during the past 2 weeks, including today" on a 5-point scale ranging from *Not at all* to *Extremely*. A mean subscale score of 3 or higher is indicative of significant psychological problems and mean item responses of diagnosable cases generally range from 2.5 to 3.5 (Watson et al., 2007, 2008). In Sample 3, the IDAS-Dysphoria had an α of .89.

Patient Health Questionnaire (PHQ-8; Spitzer, Kroenke & Williams, 1999). The PHQ-9 assesses to what extent the presence of nine depressive symptoms have bothered an individual over the last two weeks. Item 9 assesses the presence of suicidal thoughts and was not administered here. Each item is rated on a 4-point Likert scale that ranges from *Not at all* to *Nearly every day*. The scale has good psychometric properties (Kroenke et al., 2001; 2010). In Sample 3, the PHQ-8 had an α of .86.

Generalised Anxiety Disorder-7 (GAD-7; Spitzer, Kroenke, Williams & Lowe, 2006). The GAD-7 assesses to what extent the presence of seven anxiety problem symptoms have bothered an individual over the last two weeks. Each item is rated on a 4-point Likert scale that ranges from *Not at all* to *Nearly every day*. Although the GAD-7 was developed to measure Generalised Anxiety Disorder, it has been shown to provide reasonable sensitivity and specificity as a screen for panic disorder, social anxiety disorder, and posttraumatic stress disorder (Kroenke et al., 2007). The scale has strong psychometric properties (Kroenke et al., 2010; Spitzer et al., 2006). In Sample 3, the GAD-7 had an α of .89.

Life Orientation Test-Revised (LOT-R; Scheier, Carver & Bridges, 1994). The LOT-R is a 6-item scale of generalized expectations of optimism/pessimism; 3 items are framed in relation to optimism and 3 in relation to pessimism and the latter items are reverse scored. Each item is rated on a 5-point Likert scale ranging from *I disagree a lot* to *I agree a lot*. The LOT-R has good psychometric properties (Scheier et al., 1994). In Sample 3, the LOT-R had an α of .85.

5.3.4. Measures Completed by Sample 4

Inventory of Statements about Self-Injury (ISAS; Klonsky & Glenn, 2009). The ISAS includes 39 items that measure 13 functions of NSSI. Each item is rated from 0 (*Not*

relevant) to 2 (*Very relevant*) and an average score is created for each function. The scale has some published psychometric properties (Klonsky & Glenn, 2009). The ISAS' subscales demonstrated the following alpha coefficients in Sample 4: Affect Regulation (α = .73), Interpersonal Boundaries (α = .80), Self-Punishment (α = .80), Self-Care (α = .76), Anti-Dissociation (α = .83), Anti-Suicide (α = .85), Sensation-Seeking (α = .58), Peer-Bonding (α = .66), Interpersonal Influence (α = .68), Toughness (α = .79), Marking Distress (α = .71), Revenge (α = .82), Autonomy (α = .77).

Inventory of Motivations for Suicide Attempts (IMSA; Ma & Klonsky, 2013). The IMSA consists of 54 items that measure 11 functions of SA in relation to the respondent's most recent SA. Each item is rated on a 5-point Likert scale ranging from *Not at all important* to *Most important* and an average score is created for each function. The scale has some published psychometric properties (Ma & Klonsky, 2013). The IMSA's subscales demonstrated the following alpha coefficients in Sample 4: Hopelessness (α = .81), Psychache/unbearability (α = .87), Escape (α = .82), Problem solving (α = .69), Impulsivity (α = .78), Burdensomeness (α = .87), Belongingness (α = .78), Fearlessness/capability (α = .78), Influencing others (α = .89), Help seeking (α = .77). The IMSA "Other" subscale was not included in our analyses because its items were only weakly intercorrelated (α = .47).

5.3.5. Measures Completed by Sample 5

Beck Scale for Suicide Ideation (BSS; Beck & Steer, 1991). The BSS is a widely used 21-item self-report measure of passive suicide desire, active suicide desire, perceived capability to make a SA, and specific plans and preparations for a SA. The BSS contains two optional items (20 and 21) that assess previous suicide attempts; these items were not administered here as this content was assessed by questions measuring Self-Injurious Thoughts and Behavior (above). The BSS has strong psychometric properties (Beck & Steer, 1991; Brown, 2001). Although a total score is often computed, factor analytic results generally support a two factor solution where one factor (Suicidal Desire and Ideation) measures a desire for death, frequency of suicidal ideation, and lacking deterrents for suicide, and a second factor (Resolved Plans and Preparations) measures specific plans and suicidal intent (see Holden & DeLisle, 2005; Witte et al., 2006). However, as there is no consensus on the exact composition of the two factors, we determined the best fitting measurement model for our sample by conducting a Maximum Likelihood Exploratory Factor Analysis (EFA) with Promax rotation, specifying a two factor solution. Our results were similar to those observed in previous research (see Holden & DeLisle, 2005; Witte et al., 2006). In Sample 5, the Suicidal Desire and Ideation factor (items 1-9) had an α of .92 and the Resolved Plans and

Preparations factor (items 12-14 and 16-18) had an α of .75. Items 10, 11, 15, and 19 were not analysed because they demonstrated weak factor loadings (> .32) or cross-loaded (< .32 on both factors).

Suicide Cognitions Scale (SCS; Rudd et al., in preparation). The SCS is an 18 item self-report measure of suicidal beliefs. Items are rated on a 5-point Likert scale ranging from *Strongly disagree* to *Strongly agree*. Although the scale is unpublished, two studies have demonstrated that it has good psychometric properties, although there has been debate regarding whether a two- or three-factor structure is optimal (Bryan et al., 2014; Ellis & Ruffiano, 2015). A two-factor solution fit our data best (see Supplementary material for Chapter 5). The Unlovability factor (items 1, 2, 4, 6, 7, 9, 14-16, 18) had an α of .92 and the Unbearability factor (items 3, 5, 8, 11-13) had an α of .93. Items 10 and 17 were not analysed because they cross-loaded on both factors.

Interpersonal Needs Questionnaire (INQ; Van Orden et al., 2012). The INQ is a 15-item measure of the belief that other people would be better off without the respondent (Perceived burdensomeness) and a perception of a lack of interpersonal connections (Thwarted belongingness). Items are rated on a 7-point Likert scale ranging from *Not at all true for me* to *Very true for me*. The INQ has good psychometric properties (Hill et al., 2015; Van Orden et al., 2012). In Sample 5, the Perceived burdensomeness factor had an α of .93 and the Thwarted belongingness factor had an α of .89.

Mclean Screening Instrument for Borderline Personality Disorder (MSI-BPD; Zanarini et al., 2003). The MSI-BPD is a 10-item self-report measure which identifies individuals who are likely to meet the diagnostic criteria for Borderline Personality Disorder. Each DSM-IV BPD criterion is assessed with one item on the MSI-BPD, except for the paranoid ideation/dissociative symptoms criterion that is measured with two items on the MSI-BPD. Each item is rated on a *Yes/No* basis and a total score is computed. The scale has some demonstrated psychometric properties and a cutoff of 7 or more yielded good sensitivity (.81) and specificity (.85) for the diagnosis of DSM-IV BPD (Zanarini et al., 2003). Item 2 assesses intentional physical injury with and without suicidal intent; this item was omitted to avoid confounding results. In Sample 5, the MSI-BPD had an α of .74.

Brief Experiential Avoidance Questionnaire (BEAQ; Gamez et al., 2014). The BEAQ is a 15-item measure of experiential avoidance, which can be defined as an unwillingness to remain in contact with distressing emotions, thoughts, memories, and physical sensations, even when doing so creates harm in the long run (Hayes, Strosahl, & Wilson, 1999). Items are rated on a 6-point Likert scale ranging from *Strongly disagree* to

Strongly agree. The scale has reasonable psychometric properties (Gamez et al., 2014). In Sample 5, the BEAQ had an α of .86.

Short Form Perceived Stress Scale (PSS-4; Cohen et al., 1983). The PSS-4 is a 4-item self-report measure of the subjective experience of stress, rated for the past month. Respondents use a 5-point Likert scale ranging from *Never* to *Very often*. The scale has good psychometric properties (Cohen et al., 1983; Warttig et al, 2013). In Sample 5, the PSS-4 had an α of .80.

Multidimensional Scale of Perceived Social Support (MSPSS; Zimet, Dahlem, Zimet & Farley, 1988). The MSPSS is a 12-item self-report instrument designed to assess perceived social support from family, friends, and significant others. Items are rated on a 7-point Likert scale ranging from *Very strongly disagree* to *Very strongly agree* and a total score is computed (Osman et al., 2014). The scale has good psychometric properties (Zimet et al., 1990). In Sample 5, the MSPSS had an α of .90.

5.3.6. Measures Completed by Sample 6

Depressive Symptom Inventory-Suicidality Subscale (DSI-SS; Metalsky & Joiner, 1997). The DSI-SS is a 4-item self-report questionnaire designed to identify the frequency and intensity of suicidal ideation and impulses in the past two weeks, rated on a 4-point scale. The scale has been shown to have reasonable psychometric properties (Joiner, Pfaff & Acres, 2002). The DSI-SS had an α of .89 in Sample 6.

Difficulties in Emotion Regulation Scale (DERS; Gratz & Roemer, 2004). The DERS is a 36-item self-report measure of difficulties regulating various dimensions of negative emotion. Items are rated on a 5-point Likert scale ranging from *Almost never* to *Almost always*. Factor analytic studies to date have produced mixed results, although the majority of studies support a six factor model (Fowler et al., 2014). The DERS' subscales demonstrated the following alpha coefficients in Sample 6: Nonacceptance of emotional responses (α = .90), Difficulty engaging in goal-directed behavior when experiencing negative emotions (α = .86), Impulse control difficulties when experiencing negative emotions (α = .81), Lack of emotional awareness (α = .84), Limited access to emotion regulation strategies (α = .87), and Lack of emotional clarity (α = .87).

Difficulties in Emotion Regulation Scale-Positive (DERS-Positive; Weiss, Gratz & Lavender, 2015). The DERS-Positive is a 13-item self-report measure of difficulties regulating various dimensions of positive emotion. Items are rated on a 5-point Likert scale ranging from *Almost never* to *Almost always*. The DERS-Positive's subscales demonstrated the following alpha coefficients in Sample 6: Nonacceptance of Positive Emotions (α = .87),

Difficulties engaging in goal-directed behavior when experiencing positive emotions (α = .93), and Difficulties controlling behaviors when experiencing positive emotions (α = .93).

Emotion Reactivity Scale (ERS; Nock, Wedig, Holmberg & Hooley, 2008). The ERS is a 21-item self-report measure of the sensitivity, intensity, and duration of emotions, rated on a 5-point Likert scale ranging from *Not at all like me* to *Completely like me*. The scale has some demonstrated psychometric properties that were based on a single, small sample (Nock et al., 2008). The ERS had an α of .91 in Sample 6.

Brief Reasons for Living Scale (BRFLS; Ivanoff, Jang, Smyth & Linehan, 1994). The BRFLS is a short measure of reasons for living. There are six subscales and items are rated on a 6-point Likert scale ranging from *Not at all important* to *Extremely important*. The scale has good psychometric properties and is widely used (Ivanoff et al., 1994). The RFL's subscales demonstrated the following alpha coefficients in Sample 6: Survival and coping beliefs (α = .49), Responsibility to family (α = .79), Child-related concerns (α = .95), Fear of suicide (α = .70), Fear of social disapproval (α = .81), and Moral objections (α = .75).

Current mood. Participants were asked to select which of eight moods (energetic/alert, enthusiastic/euphoric, peaceful/serene, relaxed/calm, tired/sluggish, sad/down, tense/upset, anxious/jittery) best described how they "currently feel, right now." The eight affective states used were adapted from the 12-Point Affect Circumplex Scales (Yik, Russell & Steiger., 2011) and were used to compute four mutually exclusive combinations of valence (positive/negative) and arousal (activated/deactivated).

Satisfaction with Life Scale (SWLS; Diener, Emmons, Larsen & Griffin, 1985). The SWLS is a 5-item measure of participants' global assessments of how satisfied they are with their lives. Items are rated on a 7-point Likert scale ranging from *Strongly disagree* to *Strongly agree*. The SWLS had an α of .86 in Sample 6.

Subjective Happiness Scale (SHS; Lyubomirsky & Lepper, 1999). The SHS contains 4 items that are assessed on a 7-point Likert scale. 2 items ask respondents to characterize their happiness using absolute ratings and ratings relative to peers; 2 items offer brief descriptions of happy and unhappy individuals and ask respondents the extent to which each characterization describes them. Responses to the 4 items are combined. The SHS has reasonably good published psychometric properties (Lyubomirsky & Lepper, 1999). However, Item 4, which is reverse-scored, performed quite oddly in Sample 6; the distribution of responses to items 1-3 demonstrated a fairly neat positive skew, whereas the distribution of responses to item 4 resembled an inverted U, suggesting that respondents found this item unclear. Item 4 demonstrated an almost zero correlation with items 1-3 and

drastically impacted α. This item was therefore omitted from analyses; items 1-3 demonstrated an α of .90.

Subjective Vitality Scale (SVS; Ryan & Frederick, 1997). The SVS consists of 6 items and taps perceptions of being full of energy and alive rated on a 1 (*Not at all true*) to 7 (*Very true*) scale (Bostic, Rubio & Hood., 2000; Ryan & Frederick, 1997). The SVS has good psychometric properties (Bostic et al., 2000; Ryan & Frederick, 1997). The SVS had an α of .88 in Sample 6.

Missing Data

There was relatively small amounts of missing data in several samples (>8%, although the ISAS and IMSA were missing 23% and 44% of data; see Supplementary material for Chapter 5), which was not missing completely at random (MCAR). We multiply imputed missing data on all variables at the item level (Gottschall, West & Enders, 2012) using SPSS version 21.0 (IBM Corp, 2012). Predictive mean matching imputation was used because the data were somewhat skewed (White, Royston & Wood, 2011) and the number of imputations was matched to the percentage of missing information (Bodner, 2008; Graham, Olchowski & Gilreath, 2007; White et al., 2011). Multiple imputation is increasingly advocated as the optimal approach for dealing with missing data (Graham, 2009; Schafer & Graham, 2002; Shrive, Stuart, Quan & Ghali, 2006) and there is evidence that MI performs well across different circumstances, such as small samples, very large multiple regressions, and when there are large amounts of missing data (Graham & Schafer, 1999).

5.4. Results

5.4.1. Internal Consistency and Test-Retest Reliability

Table 1 presents internal consistency reliabilities (coefficient alphas) and average interitem correlations (AICs) for the SABS and NSIBS across four samples. The alpha reliabilities are consistently strong. They range from .75 to .89. Eight of the 76 coefficients fall below α = .80, and every subscale reaches or exceeds α = .80 in at least two samples. We also computed AICs because they provide a straightforward measure of internal consistency that is not affected by the number of items and because coefficient alpha underestimates true reliability (Novick & Lewis, 1967; Sijtsma, 2009). AICs of around .15 measure relatively broad constructs and AICs of around .50 measure relatively narrow constructs (Briggs & Cheek, 1986; Clark & Watson, 1995). The subscales of the SABS and NSIBS demonstrate high AICs, indicating that each subscale measures a relatively narrow and specific dimension of SICs. The AICs for the total score on each scale provides evidence that the SABS and the NSIBS tap a range of moderately related content.

Table 1 also presents the test retest reliabilities of the SABS and NSIBS over 2-4 weeks (*M* = 18 days, Median = 17 days for Samples 5 and 6). The ICC statistics indicate "moderate" to "excellent" agreement over time (Fleiss, 1986; Nunnally & Bernstein, 1994). The NSIBS subscales and total score tend to demonstrate higher test retest reliability than the SABS subscales and total score.

5.4.2. Descriptive Statistics

SABS and NSIBS Subscale Means and Standard Deviations. Table 1 presents the means and standard deviations for the SABS and NSIBS in four samples. Our samples endorsed the SABS *Stigma* (*M* = 4.79), *Self-punishment* (*M* = 4.68), and *Escape* (*M* = 4.09) subscales most strongly and the *Belonging* (*M* = 1.82) and *Revenge* (*M* = 2.37) subscales the least. They endorsed the NSIBS *Self-punishment* (*M* = 5.34), *Escape* (*M* = 5.32), and *Stigma* (*M* = 5.17) subscales most strongly and the *Belonging* (*M* = 1.93) and *Interpersonal influence* (*M* = 2.44) subscales the least. On average, the NSIBS subscales were endorsed more strongly (*M* = 4.28) than the SABS subscales (*M* = 3.44).

Relationships between demographic characteristics and the subscales of the SABS and NSIBS were explored in a series of analyses in Samples 5 and 6, our two largest samples. All post-hoc comparisons were exploratory. Tables presenting these results are displayed in the Supplementary material for Chapter 5 and the main findings are discussed below.

Table 1. Means, Internal Consistency Reliabilities (Coefficient Alphas), Average Interitem Correlations (AICs), and Test Retest Reliabilities for the Suicide Attempt Beliefs Scale (SABS) and Nonsuicidal Self-Injury Beliefs Scale (NSIBS)

	Sample 3 (*N* = 484)				Sample 4 (*N* = 380)[1]			Sample 5 (*N* = 664; 130[2])					Sample 6 (*N* = 650; 135[2])				
Subscales (Number of Items)	*M*	*SD*	α	AIC	*M*	*SD*	α	*M*	*SD*	α	AIC	ICC	*M*	*SD*	α	AIC	ICC
SABS (26)																	
Self-punishment (3)	4.83	1.75	.80	.58	4.59	1.98	.84	4.64	2.01	.87	.69	.73	4.65	2.01	.86	.68	.82
Escape (4)	4.20	1.58	.77	.46	4.17	1.80	.82	3.98	1.80	.82	.54	.67	4.01	1.75	.82	.53	.63
Dependence (5)	2.91	1.45	.81	.47	3.15	1.55	.81	2.96	1.48	.80	.44	.54	3.00	1.47	.80	.45	.63
Belonging (3)	1.88	1.13	.80	.57	1.93	1.32	.85	1.75	1.16	.84	.64	.49	1.73	1.11	.83	.64	.58
Revenge (4)	2.48	1.48	.81	.51	2.51	1.57	.81	2.27	1.46	.82	.54	.64	2.21	1.40	.81	.52	.69
Stigma (4)	5.08	1.39	.75	.43	4.55	1.70	.78	4.79	1.67	.82	.53	.67	4.73	1.73	.83	.55	.71
Eliciting help (3)	3.66	1.69	.78	.55	3.17	1.70	.77	3.29	1.71	.81	.59	.71	3.19	1.71	.81	.60	.66
Total score (range = 26-182)	92.68	24.44	.88	.22	89.74	53.15	.91	88.03	29.71	.92	.30	.70	87.48	29.36	.92	.30	.71
NSIBS (39)																	
Self-punishment (4)	5.50	1.65	.89	.67	5.08	1.72	.85	5.38	1.56	.86	.60	.68	5.39	1.57	.87	.63	.84
Dependence (4)	3.97	1.65	.81	.62	3.89	1.74	.83	4.10	1.65	.82	.53	.79	4.11	1.60	.82	.54	.81
Escape (4)	5.38	1.55	.85	.58	5.26	1.64	.87	5.37	1.50	.83	.55	.75	5.26	1.54	.83	.55	.74
Anti-dissociation (3)	5.15	1.71	.84	.63	4.78	1.78	.81	5.02	1.68	.83	.62	.73	4.91	1.72	.84	.64	.80
Problematic (5)	4.92	1.31	.76	.39	4.77	1.46	.80	4.95	1.38	.78	.42	.85	4.92	1.40	.81	.46	.83
Anti-suicide (3)	4.90	1.84	.86	.66	4.76	1.89	.86	4.92	1.85	.87	.70	.75	4.85	1.89	.89	.73	.82
Enjoyable (4)	4.09	1.67	.83	.54	3.70	1.68	.81	4.09	1.63	.81	.51	.71	4.09	1.62	.82	.53	.78
Belonging (3)	1.91	1.21	.81	.59	2.07	1.43	.83	1.90	1.26	.84	.65	.62	1.83	1.19	.83	.63	.69
Stigma (4)	5.33	1.40	.80	.51	4.83	1.61	.79	5.29	1.48	.83	.55	.76	5.21	1.50	.83	.54	.77
Interpersonal influence (5)	2.47	1.40	.84	.51	2.74	1.37	.87	2.32	1.32	.83	.51	.63	2.21	1.26	.83	.50	.82
Total score (range = 39-273)	169.91	36.28	.92	.22	162.10	78.07	.94	168.83	37.47	.93	.24	.77	166.66	36.25	.92	.22	.88

Note. ICC = Single Measure Two-Way Mixed Absolute Agreement Intraclass Correlation Coefficient of Test Retest Reliability. [1]It was not possible to compute an AIC for Sample 4 using multiply imputed data; [2]Denotes Time 2 sample sizes.

Age. Correlations between age and the subscales of the SABS and NSIBS were consistently "small" (Cohen, 1998) and negative. Correlations between age and the subscales of the SABS ranged from r = .00 to r = -.15, with most scores ~r = .07. About half of these correlations were statistically significant. *Eliciting help* demonstrated the largest correlations with age (r = -.12, -.15) and was the only SABS subscale to demonstrate a statistically significant negative correlation with age in both samples. Correlations between age and the subscales of the NSIBS ranged from r = -.03 to r = -.23 (most scores ~r = .12) and were generally statistically significant. *Self-punishment*, *Dependence*, *Anti-dissociation*, *Problematic*, *Anti-suicide*, and *Stigma* all demonstrated statistically significant negative correlations with age in both samples.

Gender. There were relatively few statistically significant main effects for gender across the subscales of the SABS and NSIBS. The SABS Stigma and NSIBS *Anti-dissociation* and *Belonging* subscales demonstrated statistically significant post-hoc differences, although there was no particular pattern to these differences, with males, for example, sometimes demonstrating higher average scores on a particular subscale and sometimes demonstrating lower average scores on another subscale. None of the SABS or NSIBS subscales demonstrated statistically significant main effects for gender across both samples.

Ethnicity. There were relatively few statistically significant main effects for ethnicity across the subscales of the SABS and NSIBS. The SABS *Stigma* subscale was the only subscale to demonstrate statistically significant post-hoc differences; people reporting a Mixed/Multiple ethnic group ethnicity, on average, scored higher than people reporting an Asian (Indian, Pakistani, Bangladeshi, Chinese) ethnicity. None of the SABS or NSIBS subscales demonstrated statistically significant main effects for ethnicity in both samples.

Marital Status. Seven of the 17 SABS/NSIBS subscales demonstrated statistically significant main effects for marital status. There was a pattern across five of the subscales (SABS *Self-punishment*, NSIBS *Anti-dissociation*, *Problematic*, *Escape*, and *Belonging*) for single people to score higher on particular subscales, on average, than married people. None of the SABS or NSIBS subscales demonstrated statistically significant main effects for marital status in both samples.

Employment Status. Statistically significant main effects were observed in about half of the SABS and NSIBS subscales for employment status. Unemployed people tended, on average, to score higher on SABS and NSIBS subscales than people who were employed, self-employed, or a student. Students tended, on average, to score higher on SABS and

NSIBS subscales than people who were employed or self-employed, although post-hoc differences were not always statistically significant. The SABS *Dependence* and NSIBS *Interpersonal influence* and *Anti-suicide* subscales demonstrated statistically significant main effects for employment status across both samples.

Level of Education. Statistically significant main effects were observed in five of the seven SABS subscales and none of the ten NSIBS subscales for level of education. Post hoc comparisons indicated that people with a Master's degree tended, on average to score statistically significantly lower on SABS and NSIBS subscales than people with higher or lower levels of education. There were no other consistent post hoc group differences across levels of education. The SABS *Stigma*, *Self-punishment*, and *Escape* subscales and NSIBS *Anti-dissociation, Dependence, Anti-suicide, Enjoyable,* and *Belonging* subscales all demonstrated statistically significant main effects for employment status in both samples.

Annual Income. Statistically significant main effects were observed for one of the seven SABS subscales and eight of the ten NSIBS subscales for annual income. Very few of the post hoc comparisons were statistically significant. There appeared to be a trend for scores to be lower in the higher income categories than the lower income categories but these post hoc comparisons were rarely statistically significant. None of the SABS or NSIBS subscales demonstrated statistically significant main effects for income in both samples.

Current Mood. We also examined whether responses to the SABS and NSIBS are influenced by current mood. Statistically significant relationships were found between current mood and the SABS *Self-punishment* ($F(3,646) = 5.81$, $p < .001$) and *Dependence* ($F(3,646) = 5.14$, $p < .01$) subscales. Posthoc comparisons indicated that, for both subscales, participants experiencing a deactivated positive current mood (peaceful/serene, relaxed/calm) reported statistically significantly lower scores on these subscales than participants experiencing an activated negative current mood (tense/upset, anxious/jittery) or a deactivated negative current mood (tired/sluggish, sad/down).

Statistically significant relationships were found between current mood and the NSIBS *Escape* ($F(3,646) = 4.49$, $p < .01$), *Self-punishment* ($F(3,646) = 6.42$, $p < .001$), *Anti-dissociation* ($F(3,646) = 3.34$, $p < .05$), *Stigma* ($F(3,646) = 3.35$, $p < .05$), *Dependence* ($F(3,646) = 9.92$, $p < .001$), and *Anti-suicide* ($F(3,646) = 2.67$, $p < .05$) subscales. Again, participants reporting a deactivated positive current mood reported statistically significantly lower scores on these subscales than participants reporting an activated or deactivated negative current mood.

Overall, these results indicate that being in a deactivated positive current mood (feeling peaceful/serene, relaxed/calm) is associated with a lower likelihood of endorsing some subscales of the SABS and NSIBS than being in an activated negative current mood (tense/upset, anxious/jittery) or a deactivated negative current mood (tired/sluggish, sad/down). Interestingly, participants reporting an activated positive current mood (energetic/alert, enthusiastic/euphoric) did not differ significantly from those reporting other moods.

5.4.3. Internal Structure of the SABS and NSIBS

Subscale Correlations. For ease of presentation, and because means, standard deviations, and correlations do not meaningfully differ for Samples 5 and 6 (participants were randomly assigned to either study during data collection), we combined both samples to examine the correlations among the subscales of the SABS and the NSIBS (see Tables 2 and 3). The SABS and NSIBS show good discriminant validity, with correlations generally in the moderate range. These data demonstrate that specific facets of SICs can be clearly distinguished from one another. The strongest correlations amongst the SABS subscales were observed between the *Revenge* and *Belonging* subscales (r = .59 and .63). The strongest correlations amongst the NSIBS subscales were observed between the *Interpersonal influence* and *Belonging* subscales (r = .62 and .70); the strength of these latter correlations indicates that these subscales are closely related, which is consistent with the fact that items from these scales merged to form one factor (*Eliciting help*) during the development of the SABS.

Table 4 presents correlations between the SABS and NSIBS. These analyses provide important information regarding how closely related cognitions about SA are to cognitions about NSSI. These analyses are especially important given that five subscales (*Self-punishment*, *Escape*, *Dependence*, *Belonging* and *Stigma*) convey such a similar theme in the SABS and NSIBS that these subscales were assigned the same label. The correlations presented in Table 4 show that the SABS is generally moderately correlated with the NSIBS. Many of the SABS subscales demonstrate virtually no correlation with the NSIBS subscales. It is particularly noteworthy that the subscales with identical labels (e.g., *Self-punishment*) exhibit "moderate" to "large" correlations (rs range from .22 to .61). Of the scales that carry identical labels, the *Self-punishment* subscales were most strongly correlated (rs = .53, .59, .61) and the *Escape* subscales were least strongly correlated (rs = .22, .35, .37). These results provide further evidence to support the discriminant validity of the subscales that make up the SABS and NSIBS.

Table 2 Suicide Attempt Beliefs Scale (SABS) Subscale Correlations

	Self-punishment	Escape	Dependence	Belonging	Stigma	Eliciting help	Revenge
Self-punishment		.56**	.49**	.13**	.51**	.28**	.10
Escape	.60**		.59**	.22**	.51**	.37**	.23**
Dependence	.55**	.61**		.40**	.36**	.35**	.37**
Belonging	.16**	.26**	.37**		.19**	.46**	.63**
Stigma	.60**	.54**	.39**	.13**		.37**	.23**
Eliciting help	.38**	.40**	.37**	.47**	.38**		.56**
Revenge	.22**	.28**	.39**	.59**	.21**	.57**	

Note. Correlations for Samples 5 and 6 combined are shown below the diagonal (N = 1,314); correlations for Sample 4 are presented above the diagonal (N = 380). ** = $p < .01$ (2-tailed).

Table 3 Nonsuicidal Self-Injury Beliefs Scale (NSIBS) Subscale Correlations

	Self-punishment	Dependence	Escape	Anti-dissociation	Problematic	Anti-suicide	Enjoyable	Belonging	Stigma	Interpersonal influence
Self-punishment		.50**	.52**	.51**	.45**	.51**	.43**	.13**	.52**	.30**
Dependence	.50**		.51**	.48**	.29**	.54**	.51**	.29**	.44**	.47**
Escape	.45**	.54**		.69**	.35**	.49**	.58**	.12**	.48**	.20**
Anti-dissociation	.42**	.46**	.59**		.30**	.47**	.56**	.17**	.45**	.23**
Problematic	.36**	.16**	.19**	.26**		.34**	.19**	.12**	.50**	.23**
Anti-suicide	.48**	.53**	.45**	.47**	.24**		.37**	.20**	.43**	.33**
Enjoyable	.31**	.45**	.63**	.50**	-.03	.29**		.32**	.39**	.33**
Belonging	.09**	.21**	.16**	.16**	-.02	.12**	.28**		.19**	.70**
Stigma	.47**	.45**	.44**	.42**	.44**	.44**	.30**	.10**		.29**
Interpersonal influence	.15**	.20**	.16**	.15**	.09**	.19**	.22**	.62**	.15**	

Note. Correlations for Samples 5 and 6 combined are shown below the diagonal (N = 1,314); correlations for Sample 4 are presented above the diagonal (N = 380). ** = p < .01 (2-tailed).

Table 4 Correlations between the Suicide Attempt Beliefs Scale (SABS) and Nonsuicidal Self-Injury Beliefs Scale (NSIBS)

	NSIBS subscales									
SABS subscales	Self-punishment	Dependence	Escape	Anti-dissociation	Problematic	Anti-suicide	Enjoyable	Belonging	Stigma	Interpersonal influence
					Sample 4 (N = 380)					
Self-punishment	.53***	.38***	.25***	.25***	.20***	.33***	.25***	.07	.33***	.14*
Escape	.19***	.30***	.22***	.20***	.13*	.25***	.21***	.14***	.18***	.18***
Dependence	.20***	.41***	.11*	.11*	.00	.24***	.16***	.22***	.16***	.26***
Belonging	.08	.16***	.01	.04	.07	.09	.12*	.51***	.13*	.43***
Stigma	.23***	.22***	.17***	.16***	.25***	.26***	.20***	.15***	.44***	.13*
Eliciting help	.17***	.22***	.09	.07	.15***	.20***	.12*	.36***	.21***	.50***
Revenge	.05	.12*	-.04	-.06	.07	.12*	.11*	.44***	.11	.50***
					Sample 5 (N = 664)					
Self-punishment	.61***	.46***	.28***	.30***	.18***	.47***	.20***	.13***	.38***	.19***
Escape	.32***	.40***	.37***	.35***	.13***	.44***	.28***	.23***	.33***	.21***
Dependence	.35***	.53***	.27***	.26***	.02	.34***	.28***	.28***	.23***	.26***
Belonging	.16***	.18***	.14***	.13***	.07	.16***	.20***	.60***	.10*	.46***
Stigma	.30***	.33***	.20***	.27***	.24***	.39***	.13***	.12***	.52***	.15***
Eliciting help	.28***	.23***	.19***	.18***	.13***	.27***	.19***	.32***	.24***	.55***
Revenge	.17***	.16***	.12***	.05	.05	.16***	.17***	.42***	.15***	.60***
					Sample 6 (N = 650)					
Self-punishment	.59**	.42**	.33**	.31**	.20**	.41**	.20**	.09*	.43**	.07
Escape	.27**	.32**	.35**	.33**	.12**	.37**	.25**	.17**	.27**	.18**
Dependence	.26**	.46**	.22**	.24**	.02	.31**	.24**	.25**	.23**	.21**
Belonging	.07	.12**	.07	.09*	.00	.11**	.13**	.57**	.04	.49**
Stigma	.25**	.26**	.21**	.25**	.25**	.29**	.10*	.07	.48**	.11**
Eliciting help	.14**	.21**	.18**	.16**	.06	.20**	.12**	.38**	.13**	.57**
Revenge	.10**	.14**	.09*	.09*	-.01	.09*	.18**	.43**	.10*	.58**

Note. * = p < .05 (2-tailed); *** = p < .001 (2-tailed).

Table 5 Correlations Between the Suicide Attempt Beliefs Scale (SABS) and Nonsuicidal Self-Injury Beliefs Scale (NSIBS) and Measures of Depressive Symptoms, Anxiety Symptoms, and Optimism

		Subscales that are apparent in the SABS and the NSIBS					Other SABS subscales		Other NSIBS Subscales				
Scale		Self-punishment	Escape	Dependence	Belonging	Stigma	Revenge	Eliciting help	Anti-dissociation	Problematic	Anti-suicide	Enjoyable	Interpersonal influence
PHQ-8	SABS	**.38*****	**.24*****	.29***	.04	**.32*****	.01	.04					
	NSIBS	.24***	.12**	**.40*****	.06	.18***			.18**	.04	.33**	.12**	.02
GAD-7	SABS	**.35*****	.23**	.26***	.01	.34***	.00	.03					
	NSIBS	.23***	.18***	**.40*****	.03	.24***			.21**	.07	.31**	.15**	.00
IDAS-Dysphoria	SABS	**.47*****	.23***	.29***	.03	.32***	.02	.06					
	NSIBS	.32***	.16***	**.44*****	.05	.23***			.21**	.06	.36**	.16**	.04
LOT-R	SABS	**-.45*****	-.19***	-.24***	.02	-.23***	-.04	-.04					
	NSIBS	-.31***	-.19***	**-.40*****	-.02	-.23***			-.19**	-.02	-.33**	-.15**	-.05

Note. Correlations in bold are statistically significantly larger than the corresponding coefficient for the similarly themed SABS or NSIBS subscale. * = $p < .05$ (2-tailed); *** = $p < .001$ (2-tailed). PHQ-8 = Patient Health Questionnaire; GAD-7 = Generalized Anxiety Disorder-7; IDAS-Dysphoria = Inventory of Depression and Anxiety Symptoms – Dysphoria subscale; LOT-R = Life Orientation Test-Revised.

5.4.4. Convergent and Discriminant Validity

We now examine the convergent and discriminant validity of the SABS and NSIBS in relation to a range of other clinical and well-being measures (see Tables 5 to 8). Each Table is presented so as to allow easy comparison of the strength of correlations between the identically labelled SABS and NSIBS subscales. For example, comparing the strength of the correlation between the SABS *Self-punishment* subscale and the PHQ-8 with the strength of the NSIBS *Self-punishment* subscale and the PHQ-8. We tested whether subscales that are apparent in the SABS and the NSIBS had statistically significantly different correlations with other variables using Steiger's (1980) modification of the Hotelling test for two correlations involving a common variable.

Correlations with Depressive Symptoms, Anxiety Problems, and Optimism. Table 5 provides evidence that the SABS and NSIBS demonstrate negligible (e.g., $r = .01$) to large (e.g., $r = .47$) positive correlations with several measures of depressive symptoms and anxiety problems. Most of these correlations were "small" to "moderate" in size (Cohen, 1998). The *Self-punishment* and *Dependence* subscales generally demonstrated the strongest correlations with depressive symptoms and anxiety problems. The *Belonging, Revenge, Eliciting help, Problematic,* and *Interpersonal influence* subscales demonstrated the weakest correlations with depressive symptoms. The SABS and NSIBS were negatively correlated with optimism and correlations were generally "small" to "moderate" in size (Cohen, 1998). Overall, Table 5 provides evidence to support the convergent and divergent validity of the SABS and NSIBS.

Correlations with Existing Measures of the Reasons for SA and NSSI. Table 6 presents the correlations between the SABS and NSIBS and the IMSA and the ISAS. The most obvious point to note is that the SABS and NSIBS and IMSA and ISAS are generally only weakly associated. Many of the correlations are trivial ($\leq r = .05$). As would be expected, the SABS subscales were often more strongly associated with the IMSA subscales than the NSIBS subscales and the opposite was apparent for the NSIBS subscales in relation to the ISAS. These results indicate that measures pertaining to either SA or NSSI are somewhat more closely associated. However, these correlations were generally only modest in size.

Some of the IMSA and ISAS subscales carry the same labels (*Self-punishment, Escape, Anti-dissociation, Anti-suicide, Interpersonal influence, Revenge*) or very similar labels (*Belongingness, Influencing others*) as SABS and NSIBS subscales and might therefore be thought to measure the same constructs. It is therefore noteworthy that the strongest correlation observed between identically named subscales was only $r = .55$. A

correlation of this magnitude is some distance away from indicating equivalence (Clark & Watson, 1995). Overall, the results presented in Table 6 indicate that the SABS and NSIBS are only somewhat related to existing measures that might be assumed to tap identical content.

Correlations with Measures of Current Suicidal Ideation. Tables 7 and 8 present correlations between the SABS and NSIBS and the BSS and DSI-SS, two measures of current suicidal thinking. "Small" to "moderate" correlations (Cohen, 1998) were generally observed between the SABS and NSIBS and current suicidal thinking. The *Self-punishment* and *Dependence* subscales demonstrated "large" correlations (Cohen, 1998). When identically named SABS/NSIBS subscales were compared, SABS subscales tended to be more strongly positively correlated with current suicidal thinking than NSIBS subscales, although the difference between correlations tended to be fairly small. *Eliciting help* and *Problematic* were the only SABS/NSIBS subscales to demonstrate a stronger relationship with the BSS Plans and preparations factor than the Desire and ideation factor. These results indicate that the SABS and NSIBS were somewhat related to current suicidal thinking, particularly suicidal desire and ideation. As would be expected, the SABS was more closely related to current suicidal thinking than the NSIBS.

Table 6 Correlations Between the Suicide Attempt Beliefs Scale (SABS) and Nonsuicidal Self-Injury Beliefs Scale (NSIBS) and Existing Measures of the Reasons for Attempting Suicide and Engaging in Nonsuicidal Self-Injury

		Subscales that are apparent in the SABS and the NSIBS					Other SABS subscales		Other NSIBS Subscales				
Scale/Subscale		Self-punishment	Escape	Dependence	Belonging	Stigma	Revenge	Eliciting help	Anti-dissociation	Problematic	Anti-suicide	Enjoyable	Interpersonal influence
IMSA													
Hopelessness	SABS	.25***	.20***	.11	**-.16***	.25***	-.15*	.01					
	NSIBS	.14*	.19***	.07	-.03	.15*			.16***	.05	.14*	.10	-.10
Psychache/ Unbearability	SABS	**.26***	.26***	.12	-.18***	**.25***	-.18***	.05					
	NSIBS	.11***	.18***	.11	-.06	.11			.18***	.03	.14*	.11	-.08
Escape	SABS	**.36***	.21***	.15*	-.11	.21***	-.07	.05					
	NSIBS	.30***	.23***	.16***	-.02	.21***			.19***	.11	.17***	.13*	.01
Problem solving	SABS	.07	.07	.14*	.10	.11	.15*	.10					
	NSIBS	.04	.05	.04	.11	.09			.00	-.01	.01	.07	.09
Impulsivity	SABS	-.15*	-.10	-.01	.17***	-.09	.25***	.08					
	NSIBS	-.02	-.01	-.01	.12	.00			.01	.01	-.02	.03	.18***
Burdensome-ness	SABS	.30***	.12	.13*	-.07	.19***	-.07	-.01					
	NSIBS	.20***	.15*	.14*	-.04	.21***			.10	.06	.15*	.09	.00
Belonging-ness	SABS	.15*	.06	.13	.04	.12	.08	.04					
	NSIBS	.16***	.11	.10	.07	.18***			.08	.03	.12	.07	.06
Fearlessness/ Capability	SABS	.11	.13	.13	-.01	.14	.02	.03					
	NSIBS	.08	.09	.18***	.04	.14*			.13*	.09	.13*	.09	.07
Influencing others	SABS	**-.23***	-.20***	-.02	.29***	-.20***	.39***	.14*					
	NSIBS	-.07	-.12*	-.04	.23***	-.05			-.15***	-.01	-.07	-.03	.24***
Help seeking	SABS	-.08	-.05	.06	.24***	-.12	.25***	.30***					
	NSIBS	.01	-.04	.05	.21***	.00			-.05	.01	.00	.02	.34***
ISAS													
Affect regulation	SABS	.23***	.17***	.08	-.10	.07	-.14*	-.03					
	NSIBS	.25***	**.47***	**.27***	-.07	**.25***			.39***	.11	.23***	.31***	-.05
Interpersonal boundaries	SABS	.07	.08	.16***	.25***	.09	.27***	.15*					
	NSIBS	.02	-.12	.09	.29***	.00			-.03	-.06	.00	.13	.24***
Self-punishment	SABS	.38***	.16***	.08	-.03	.18***	-.05	.09					
	NSIBS	**.54***	.26***	**.25***	-.07	.28***			.26***	.18***	.23***	.22***	.04
Self-care	SABS	.11	.11	.07	.12*	.08	.10	.11					
	NSIBS	.15*	.15*	**.21***	.16***	.17***			.20***	.03	.14*	.22***	.20***
Anti-dissociation	SABS	.27***	.27***	.12*	-.01	.18***	-.05	.05					
	NSIBS	.25***	.35***	**.29***	.02	.24***			.55***	.11	.26***	.29***	.07
Anti-suicide	SABS	.26***	.26***	.21***	-.01	.23***	-.01	.13*					
	NSIBS	.22***	.29***	.32***	.04	.21***			.32***	.16***	.61***	.19***	.06
	SABS	-.02	.05	.13*	.22***	.01	.30***	.13*					

Sensation-seeking	NSIBS	-.05	-.11	.07	.28***	-.04			-.02	-.08	-.02	.18***	.23***
Peer-bonding	SABS	-.05	-.04	.01	.20***	-.01	.23***	.10					
	NSIBS	-.07	**-.19***	-.03	.22***	-.06			-.14*	-.10	-.04	.05	.17*
Interpersonal influence	SABS	-.07	-.01	.06	.20***	-.03	.31***	.31***					
	NSIBS	.00	-.12	.05	**.34***	.02			-.07	-.02	.08	.03	.47***
Toughness	SABS	.10	**.13***	.15*	.19***	.13*	.30***	.18***					
	NSIBS	.10	-.03	.15***	.22***	.07			.06	-.01	.13*	.25***	.24***
Marking distress	SABS	.20***	.17***	.17***	.14*	.12*	.23***	.30***					
	NSIBS	.23***	.10	.23***	.13*	**.25***			.15*	.02	.23***	.19***	.25***
Revenge	SABS	-.06	-.01	.07	.26***	-.03	.40***	.18***					
	NSIBS	-.02	**-.17***	.04	.34***	-.04			-.14*	-.06	-.02	.01	.42***
Autonomy	SABS	.04	.12	.16***	.21***	.08	.25***	.12*					
	NSIBS	.01	-.10	.05	.20***	-.04			.01	-.10	.02	.15*	.15*

Note. Correlations in bold are statistically significantly larger than the corresponding coefficient for the similarly themed SABS or NSIBS subscale. * = $p < .05$ (2-tailed); *** = $p < .001$ (2-tailed). IMSA = Inventory of Motivations for Suicide Attempts; ISAS = Inventory of Statements about Self-Injury.

Table 7 Correlations Between the Suicide Attempt Beliefs Scale (SABS) and Nonsuicidal Self-Injury Beliefs Scale (NSIBS) and Clinical and Well-Being Variables in Sample 5

		Subscales that are apparent in the SABS and the NSIBS					Other SABS subscales		Other NSIBS Subscales				
Scale/Subscale		Self-punishment	Escape	Dependence	Belonging	Stigma	Revenge	Eliciting help	Anti-dissociation	Problematic	Anti-suicide	Enjoyable	Interpersonal influence
BSS													
Desire and ideation	SABS	**.46****	**.32****	.48**	.04	**.27****	.07	.08*					
	NSIBS	.30**	.16**	.49**	.04	.18**			.14**	-.03	.36**	.13**	.03
Plans and preparations	SABS	**.33****	**.27****	.31**	.00	**.25****	.02	.09*					
	NSIBS	.13**	.07	.33**	.02	.10*			.07	-.08*	.27**	.11**	-.01
SCS													
Unlovability	SABS	**.50****	**.33****	.48**	.10*	.27**	.10*	.13**					
	NSIBS	.42**	.21**	.52**	.10*	.24**			.20**	.05	.34**	.16**	.11**
Unbearability	SABS	**.40****	**.32****	.40**	.09*	**.31****	.10*	.15**					
	NSIBS	.27**	.20**	**.48****	.07	.24**			.22**	.12**	.32**	.11**	.12**
INQ													
Perceived burdensomeness	SABS	**.48****	**.28****	.42**	.07	.28**	.11**	.10**					
	NSIBS	.38**	.19**	.49**	.08*	.24**			.24**	.09*	.34**	.14**	.07
Thwarted belongingness	SABS	.28**	.18**	.32**	-.01	.21**	.10**	.04					
	NSIBS	.23**	.10*	.38**	.03	.17**			.11**	.02	.22**	.10**	.01
MSI-BPD	SABS	.25**	.21**	.28**	.15**	.22**	.21**	.14**					
	NSIBS	.26**	.21**	.35**	.16**	.23**			.26**	.14**	.21**	.19**	.21**
BEAQ	SABS	.35**	.29**	.33**	.09*	.24**	.13**	.16**					
	NSIBS	.34**	.28**	.41**	.05	.24**			.29**	.13**	.27**	.16**	.14**
PSS-4	SABS	.00	.02	-.03	.04	.05	.01	-.01					
	NSIBS	.04	.01	.02	-.04	.07			.04	.12**	.03	.03	.00
MSPSS	SABS	-.18**	**-.10****	-.25**	.03	-.12**	-.04	.04					
	NSIBS	-.08*	.03	-.20**	.03	-.05			.01	.05	-.09*	.00	.08*

Note. Correlations in bold are statistically significantly larger than the corresponding coefficient for the similarly themed SABS or NSIBS subscale. * = $p < .05$ (2-tailed); ** = $p < .01$ (2-tailed). Desire and ideation = Beck Scale for Suicide Ideation - Suicidal Desire and Ideation subscale; Plans and preparations = Beck Scale for Suicide Ideation - Resolved Plans and Preparations subscale; Unlovability = Suicide Cognitions Scale – Unlovability subscale; Unbearability = Suicide Cognitions Scale – Unbearability subscale; Perceived burdensomeness = Interpersonal Needs Questionnaire - Perceived burdensomeness subscale; Thwarted belongingness = Interpersonal Needs Questionnaire -

Thwarted belongingness subscale; MSI-BPD = Mclean Screening Instrument for Borderline Personality Disorder; BEAQ = Brief Experiential Avoidance Questionnaire; PSS-4 = Short Form Perceived Stress Scale; MSPSS = Multidimensional Scale of Perceived Social Support.

Table 8 Correlations Between the Suicide Attempt Beliefs Scale (SABS) and Nonsuicidal Self-Injury Beliefs Scale (NSIBS) and Clinical and Well-Being Variables in Sample 6

		Subscales that are apparent in the SABS and the NSIBS					Other SABS subscales		Other NSIBS Subscales				
Scale/Subscale		Self-punishment	Escape	Dependence	Belonging	Stigma	Revenge	Eliciting help	Anti-dissociation	Problematic	Anti-suicide	Enjoyable	Interpersonal influence
DERS													
Nonacceptance	SABS	.30***	.12***	.09*	-.03	.21***	-.04	-.02					
	NSIBS	.38***	.19***	**.29*** **	-.04	**.32*** **			.29***	.28***	.27***	.10*	-.04
Goals	SABS	.10*	.08	.08*	.03	.13***	.05	.12***					
	NSIBS	**.20*** **	.13***	.19***	-.03	.18***			.13***	.18***	.16***	.03	.05
Impulse	SABS	.30***	**.27*** **	**.30*** **	.08*	.27***	.11***	.15***					
	NSIBS	.32***	.24***	.36***	.06	.29***			.28**	.22***	.29***	.15***	.12***
Awareness	SABS	.22***	.08	.14***	.00	.16***	.01	.03					
	NSIBS	.17***	.16***	.23***	.01	.13***			.13***	.06	.08*	.18**	-.01
Strategies	SABS	.35***	.22***	.27***	.00	.25***	.05	.12***					
	NSIBS	.41***	.27***	**.42*** **	.01	.28***			.32***	.21***	.33***	.18***	.05
Clarity	SABS	.24***	.17***	.18***	.02	.25***	.04	.12***					
	NSIBS	.19***	.21***	**.30*** **	.04	.22***			.28***	.17***	.16***	.21***	.03
DERS-P													
NonacceptanceP	SABS	.37***	.28***	.27***	.08*	.25***	.08*	.10*					
	NSIBS	.32***	.23***	**.39*** **	.06	.25***			.25***	.11***	.26***	.17***	.00
GoalsP	SABS	.11***	.15***	.18***	.06	.13***	.07	.07					
	NSIBS	.09*	.13***	.17***	.07	.11***			.13***	.05	.16***	.10*	.06
ImpulseP	SABS	.19***	.24***	.21**	.11***	.17***	.10*	.10*					
	NSIBS	.16***	.20***	.24***	.11***	.12***			.20***	.08	.20***	.17***	.08
BRFLS													
Survival	SABS	**-.18*** **	-.09*	-.22***	.14***	-.12***	.08*	.05					
	NSIBS	-.06	-.09*	-.26***	.07	-.10***			-.04	.06	-.12***	-.12***	.12***
Responsibility	SABS	-.08*	-.07	-.17***	-.03	-.03	-.11***	-.03					
	NSIBS	.00	.03	-.10*	-.07	.01			-.02	.12***	.01	-.09*	-.02
Children	SABS	-.05	-.07	-.05	-.01	-.01	-.01	-.09*					
	NSIBS	-.02	-.01	-.09*	.02	-.01			-.03	.00	-.04	-.03	.01
Fear	SABS	**-.14*** **	-.11***	-.12***	.09*	-.13***	.14***	.14***					
	NSIBS	-.02	-.04	-.10***	.09*	-.06			-.05	.05	-.08*	-.05	.16***
Stigma	SABS	-.04	-.03	.00	.14***	.03	.08*	.04					
	NSIBS	**.11*** **	.00	-.01	.10*	.06			.05	.16**	.01	.02	.11***
Morality	SABS	**-.18*** **	-.09*	-.22***	.14***	-.12***	.08*	.05					
	NSIBS	-.06	-.09*	-.26***	.07	-.10***			-.04	.06	-.12***	-.12***	.12***
DSI-SS	SABS	**.34*** **	**.22*** **	.38***	.00	**.29*** **	.02	.12***					

ERS	NSIBS	.21***	.13***	.40***	.05	.19***			.16***	.07	.33***	.09*	.02
	SABS	.26***	.19**	.19***	.09*	.22***	.12**	.16***					
SWLS	NSIBS	.33***	.23***	**.35*** **	.08*	.26***			.26***	.17***	.26***	.16***	.15***
	SABS	-.26***	-.18***	-.23***	.01	-.17***	.01	-.01					
SHS	NSIBS	-.18***	-.14***	-.24***	-.01	-.16***			-.12***	-.02	-.17***	-.07	.02
	SABS	-.31***	-.13***	-.20***	.05	-.22***	.04	-.07					
SVS	NSIBS	-.28***	-.16***	-.37***	.03	-.22***			-.15***	-.07	-.18***	-.09*	.06
	SABS	-.26***	-.11***	-.20***	.06	-.19***	.04	-.05					
	NSIBS	-.20***	-.16***	**-.33*** **	.03	-.19***			-.13***	-.01	-.15***	-.07	.05

Note. Correlations in bold are statistically significantly larger than the corresponding coefficient for the similarly themed SABS or NSIBS subscale. * = $p < .05$ (2-tailed); *** = $p < .001$ (2-tailed). Nonacceptance = Difficulties in Emotion Regulation Scale (DERS) - Nonacceptance of emotional responses subscale; Goals = DERS - Difficulty engaging in goal-directed behavior when experiencing negative emotions subscale; Impulse = DERS - Impulse control difficulties when experiencing negative emotions subscale; Awareness = DERS - Lack of emotional awareness subscale; Strategies = DERS - Limited access to emotion regulation strategies subscale; Clarity = DERS - Lack of emotional clarity subscale; NonacceptanceP = Difficulties in Emotion Regulation Scale-Positive (DERS-P) - Nonacceptance of positive emotions subscale; GoalsP = DERS-P - Difficulties engaging in goal-directed behavior subscale; ImpulseP = DERS-P – Impulse control difficulties subscale; Survival = Brief Reasons for Living Scale (BRFLS) - Survival and coping beliefs subscale; Responsibility = BRFLS - Responsibility to family subscale; Children = BRFLS - Child-related concerns subscale; Fear = BRFLS - Fear of suicide subscale; Stigma = BRFLS - Fear of social disapproval subscale; Morality = BRFLS - Moral objections subscale; DSI-SS = Depressive Symptom Inventory-Suicidality Subscale; ERS = Emotion Reactivity Scale; SWLS = Satisfaction with Life Scale; SHS = Subjective Happiness scale; SVS = Subjective Vitality Scale.

Correlations with Existing Measures of Suicidal Cognitions. Tables 7 and 8 present correlations between the SABS and NSIBS and the SCS, INQ, and BRFLS, which measure perceptions that are thought to underlie suicidal thoughts and/or SA (Beck et al. 1974; Rudd et al., in submission; Van Orden et al., 2008), and, in the case of the BRFLS, reasons to live. The SABS and NSIBS subscales were generally moderately positively correlated with perceptions of unlovability and unbearability, burdensomeness, and thwarted belongingness and neither scale evidenced consistently stronger associations with the SCS or INQ. The *Self-punishment* and *Dependence* subscales demonstrated stronger positive correlations with the SCS and INQ, a few of which were "large" correlations (Cohen, 1998). These results indicate that the SABS and NSIBS measure content that is not currently tapped by existing measures of suicidal cognitions.

The SABS and NSIBS generally evidenced "small" negative correlations (Cohen, 1998) with the BRFLS and the subscales of the BRFLS tended to be more strongly associated with the SABS than the NSIBS. These results are surprising, as they indicate that beliefs about SA and NSSI are barely correlated with reasons to live.

Correlations with Related Clinical Constructs. Tables 7 and 8 present correlations between the SABS and NSIBS and symptoms of BPD, experiential avoidance, perceived stress, difficulties in regulation positive and negative emotions, and emotional reactivity. These associations were generally "small" to "moderate" in magnitude (Cohen, 1998) and the SABS and NSIBS generally evidenced equally strong associations with other variables. Interestingly, the SABS and NSIBS are almost completely uncorrelated with perceived stress.

Correlations with Measures of Well-Being. Tables 7 and 8 present correlations between the SABS and NSIBS and several measures of well-being, including perceived social support (Table 7), satisfaction with life, and subjective happiness and vitality (Table 8). Again, these associations were generally "small" to "moderate" in magnitude (Cohen, 1998).

5.4.5. Group Validity

We explored whether the SABS and NSIBS could discriminate between theoretically relevant subgroups in a series of analyses that are reported in Table 9.

History of Self-Injurious Thoughts and Behavior. Cognitive theory predicts that the strength of SA cognitions is associated with greater risk of SA (e.g., Rudd, 2000, 2006) and some previous research suggests that multiple attempters should be treated as distinct from single attempters and individuals with ideation but no previous attempts (Rudd, Joiner, & Rajab, 1996). We therefore explored whether the SABS could discriminate between four groups: (1) Individuals who had never experienced suicidal thoughts, (2) individuals who had

experienced suicidal thoughts but not taken action, (3) individuals who had made one SA, and (4) individuals who had made two or more SAs (see Table 9). Across two samples, we found statistically significant main effects for five of the seven SABS subscales and a fairly clear trend of participant's increasingly endorsing the subscales of the SABS the more they had moved along the continuum from thoughts to making multiple SAs. The *Belonging* and *Revenge* subscales did not evidence statistically significant main effects.

We are not aware of a cognitive theory that links NSSI cognitions to NSSI behavior in the manner specified by cognitive theories of SA but we see no reason to suspect that greater endorsement of NSSI cognitions would not be associated with an increased risk of NSSI. We explored whether the NSIBS could discriminate between three groups: (1) Individuals who had never experienced NSSI thoughts, (2) individuals who had experienced NSSI thoughts but not taken action, (3) individuals who had engaged in NSSI on two or more occasions (Table 9). Very few people (Sample 5: $N = 1$; Sample 6: $N = 6$) had engaged in NSSI once, so this category was therefore omitted. Across two samples, we found statistically significant main effects for eight of the ten NSIBS subscales. Again, there was a fairly clear trend that participants' increasingly endorsed the subscales of the NSIBS the more they had moved along the continuum from thoughts to engaging in NSSI multiple times. The *Belonging* and *Interpersonal influence* subscales did not evidence statistically significant main effects in Sample 5 and, interestingly, participants' endorsed the *Belonging* subscale less as they progressed along the continuum towards multiple NSSI episodes. Overall, these results indicate that SA and NSSI beliefs, as measured by the SABS and NSIBS, are closely associated with SA and NSSI behavior for most of the subscales of the SABS and NSIBS.

Recency of Self-Injurious Behavior. If the SABS and NSIBS are important drivers of SA and NSSI behavior, more recent SA and NSSI behavior is likely to be associated with increased endorsement of the SABS and NSIBS. We tested this prediction by examining whether the SABS and NSIBS could discriminate between four groups who had engaged in SA or NSSI in the (1) past month, (2) past year, (3) 1-2 years ago, or (4) 2+ years ago (Table 9). We found statistically significant main effects for three of the seven SABS subscales in both samples (*Self-punishment, Escape, Dependence*). Participants' who made a SA 2+ years ago demonstrated statistically significantly lower scores on these subscales and there was a general, although not always statistically significant, trend for participants' to endorse these subscales less with the passage of time.

All of the NSIBS subscales demonstrated at least one statistically significant main effect for recency in one of the two samples. Seven of the ten NSIBS subscales demonstrated

a statistically significant main effect for recency in both samples. Post hoc comparisons indicated a fairly consistent trend over time wherein each increment of time was associated with statistically significantly lower NSIBS scores. Overall, these results indicate that people endorse the NSIBS more strongly the more recently they engaged in NSSI and people endorse some subscales of the SABS (especially *Self-punishment, Escape, Dependence*) more strongly the more recently they engaged in SA.

Sensitivity analyses. We conducted a series of sensitivity analyses, re-running these analyses whilst controlling for NSSI history in the SA analyses and SA history in the NSSI analyses. All results for history and recency of self-injurious thoughts and behavior were unchanged.

Probable BPD diagnosis. We explored the association between a likely BPD diagnosis and scores on the SABS and NSIBS (Table 9) because SA and NSSI are often assumed to be synonymous with BPD and both behaviors are symptoms of a BPD diagnosis. With one exception (NSIBS *Problematic* subscale), we found that the subscales of the SABS and NSIBS were statistically significantly more strongly endorsed by individuals who exceed the cutoff for a likely diagnosis of BPD (7/10 on the MSI-BPD) than those who scored under this cutoff. (As discussed in the Measures section, one item that assesses intentional physical injury with and without suicidal intent was omitted to avoid confounding results, making the cutoff of 7 more stringent).

Table 9 Self-Injurious Thoughts and Behavior History and Recency Group Differences for the Suicide Attempt Beliefs Scale (SABS) and Nonsuicidal Self-Injury Beliefs Scale (NSIBS)

	History of Self-Injurious Thoughts and Behavior							Recency of Self-Injurious Behavior						Probable BPD diagnosis		
	Sample 5 (N = 664)				Sample 6 (N = 650)			Sample 5 (N = 427)			Sample 6 (N = 650)			Sample 5 (N = 664)		
Subscale	Group	M	SD	F-statistic	M	SD	F-statistic	Group	M	SD	F-statistic	M	SD	F-statistic	Group	M	SD
SABS				52.60***			50.92***				9.93***			5.57***			
Self-punishment	None	2.07[a,b,f]	1.58		2.26[a,b,f]	1.74		Past month	5.72[a]	1.51		5.83[a]	1.62		Under cutoff	3.78	2.06
	Thoughts	3.64[c,d,f]	1.99		3.71[c,d,f]	2.09		Past year	5.52[b]	1.69		5.56[b]	1.50				
	Attempt	4.66[a,c,e]	1.96		4.63[a,c,e]	1.80		1-2 years ago	5.74[c]	1.46		5.22	1.66		Over cutoff	4.82***	1.96
	Multiple	5.39[b,d,e]	1.66		5.43[b,d,e]	1.60		2+ years ago	4.74[a,b,c]	1.82		4.91 [a,b]	1.74				
				36.36***			37.94***				4.28*			9.31***			
Escape	None	2.08[a,b]	1.65		2.19[a,b]	1.62		Past month	4.70[a]	1.43		5.18[a]	1.39		Under cutoff	3.35	1.88
	Thoughts	3.21[c,d]	1.76		3.28[c,d]	1.75		Past year	4.60	1.68		4.72[b]	1.52				
	Attempt	3.97[a,c,e]	1.72		3.92[a,c,e]	1.59		1-2 years ago	4.82[b]	1.48		4.62[c]	1.45		Over cutoff	4.11***	1.75
	Multiple	4.56[b,d,e]	1.59		4.62[b,d,e]	1.53		2+ years ago	4.13[a,b]	1.67		4.01[a,b,c]	1.59				
				20.54***			21.21****				14.76***			25.22***			
Dependence	None	1.77[a,b]	1.12		1.83[a]	1.24		Past month	4.06[a,b]	1.42		4.44[a,b,c]	1.36		Under cutoff	2.31	1.23
	Thoughts	2.49[c]	1.35		2.60[b]	1.34		Past year	3.42[a,c]	1.51		3.68[a,d,e]	1.44				
	Attempt	2.82[a,d,e]	1.34		2.66[c]	1.45		1-2 years ago	3.42[d]	1.50		3.16[b,d,f]	1.26		Over cutoff	3.10***	1.49
	Multiple	3.35[b,c,e]	1.49		3.41[a,b,c]	1.44		2+ years ago	2.78[b,c,d]	1.29		2.66[c,e, f]	1.33				
				1.55			1.06				1.60			4.19*			
Belonging	None	1.73	1.10		1.57	1.14		Past month	1.89	1.29		1.89	1.29		Under cutoff	1.48	.95
	Thoughts	1.88	1.25		1.83	1.22		Past year	1.73	1.17		1.84[a]	1.11				
	Attempt	1.79	1.09		1.63	0.98		1-2 years ago	1.76	1.17		1.76	1.13		Over cutoff	1.81***	1.19
	Multiple	1.67	1.12		1.70	1.07		2+ years ago	1.57	0.97		1.45[a]	0.83				
				1.89			1.35				1.09			2.39			
Revenge	None	1.78	1.17		1.92	1.46		Past month	2.46	1.51		2.47	1.45		Under cutoff	1.89	1.28
	Thoughts	2.37	1.51		2.31	1.50		Past year	2.33	1.60		2.17	1.32				
	Attempt	2.47	1.40		1.98	1.23		1-2 years ago	2.13	1.34		2.34	1.38		Over cutoff	2.35***	1.48
	Multiple	2.18	1.45		2.21	1.37		2+ years ago	2.13	1.33		1.97	1.29				
				41.23***			64.68***				1.32			2.21			
Stigma	None	2.37[a,b,e]	1.79		2.33[a,b,f]	1.91		Past month	5.37	1.18		5.51	1.22		Under cutoff	4.17	1.83
	Thoughts	4.08[c,d,e]	1.89		3.84[c,d,f]	1.84		Past year	5.37	1.30		5.51	1.19				
	Attempt	5.07[a,c]	1.20		4.80[a,c,e]	1.47		1-2 years ago	5.28	1.20		5.25	1.36		Over cutoff	4.92***	1.61

	Multiple	5.29[b,d]	1.33		5.44[b,d,e]	1.24		2+ years ago	5.10	1.38		5.14	1.39			
				5.23*			7.07***				1.42			3.78*		
Eliciting help	None	2.04[a,b]	1.44		2.09[a]	1.70		Past month	3.63	1.83		3.71	1.51		Under cutoff	2.92	1.74
	Thoughts	3.09	1.69		2.90[b]	1.75		Past year	3.62	1.69		3.53	1.64				
	Attempt	3.66[a]	1.77		3.28	1.48		1-2 years ago	3.37	1.68		3.67[a]	1.71		Over cutoff	3.37*	1.70
	Multiple	3.38[b]	1.70		3.43[a,b]	1.68		2+ years ago	3.26	1.69		3.06[a]	1.60				
	NSIBS			49.24***			30.90***				12.92***			8.77***		
Self-punishment	None	3.11[a]	2.09		3.14[a,b]	1.65		Past month	5.79[a,b,c]	1.29		5.74[a]	1.36		Under cutoff	4.86	1.85
	Thoughts	3.29[b]	1.53		4.72[a]	1.91		Past year	5.38[a,d]	1.43		5.44[b]	1.48				
	Multiple	5.52[a,b]	1.42		5.51[b]	1.48		1-2 years ago	5.23[b]	1.41		5.17	1.72		Over cutoff	5.49***	1.60
								2+ years ago	4.68[c,d]	1.73		4.81[a,b]	1.67				
				18.98***			15.83***				62.11***			67.00***		
Dependence	None	2.52[a]	1.82		2.66[a]	1.57		Past month	4.83[a,b,c]	1.39		4.82[a,b,c]	1.29		Under cutoff	3.24	1.65
	Thoughts	2.71[b]	1.60		3.13[b]	1.54		Past year	3.87[a,d,e]	1.45		3.99[a,d,e]	1.50				
	Multiple	4.20[a,b]	1.60		4.21[a,b]	1.57		1-2 years ago	2.86[b,d]	1.35		3.28[b,d,f]	1.39		Over cutoff	4.29***	1.72
								2+ years ago	2.89[c,e]	1.45		2.48[c,e,f]	1.30				
				56.64***			39.44***				12.68***			13.06***		
Escape	None	3.19[a]	2.18		3.00[a]	1.63		Past month	5.77[a,b,c]	1.18		5.62[a]	1.28		Under cutoff	4.82	1.78
	Thoughts	2.96[b]	1.85		4.01[b]	1.66		Past year	5.43[a]	1.34		5.41[b]	1.49				
	Multiple	5.52[a,b]	1.33		5.39[a,b]	1.44		1-2 years ago	4.98[b]	1.33		5.01	1.53		Over cutoff	5.49***	1.53
								2+ years ago	4.88[c]	1.74		4.51[a,b]	1.62				
				29.72***			20.33***				9.45***			9.25***		
Anti-dissociation	None	3.03[a]	2.20		2.89[a,b]	1.52		Past month	5.40[a,b]	1.43		5.31[a,b]	1.51		Under cutoff	4.35	1.96
	Thoughts	3.31[b]	1.94		4.26[a]	1.77		Past year	5.05	1.60		4.83[a]	1.75				
	Multiple	5.15[a,b]	1.58		5.01[b]	1.67		1-2 years ago	4.67[a]	1.61		4.63	1.76		Over cutoff	5.17***	1.72
								2+ years ago	4.43[b]	1.94		4.31[b]	1.84				
				25.46***			17.23***				2.69*			.51		
Problematic	None	3.24[a]	2.22		3.47[a]	1.71		Past month	5.06	1.26		5.04	1.30		Under cutoff	4.76	1.53
	Thoughts	4.11	1.95		4.26	1.73		Past year	4.83	1.30		4.94	1.35				
	Multiple	5.04[a]	1.27		5.01[a]	1.34		1-2 years ago	5.19	1.19		4.86	1.73		Over cutoff	4.99	1.46
								2+ years ago	5.32	1.29		5.10	1.19				
				34.19***			16.03***				12.35***			6.84***		
Anti-suicide	None	2.36[a]	1.82		2.99[a]	1.64		Past month	5.39[a,b,c]	1.61		5.19[a]	1.75		Under cutoff	4.39	2.25
	Thoughts	3.48[b]	1.86		3.91	1.98		Past year	4.88[a,d]	1.78		4.91[b]	1.82				
	Multiple	5.06[a,b]	1.76		4.96[a]	1.85		1-2 years ago	4.65[b]	1.85		4.74	2.02		Over cutoff	5.04***	1.89

							2+ years ago	4.08[c,d]	1.94		4.11[a,b]	2.03			
				16.78***			10.34***			7.02***			7.68***			
Enjoyable	None	2.71[a]	1.93		2.79[a]	1.45	Past month	4.36[a,c]	1.53		4.31[a]	1.56		Under cutoff	3.69	1.78
	Thoughts	2.63[b]	1.91		3.44	1.80	Past year	4.23[b]	1.57		4.20[b]	1.61				
	Multiple	4.18[a,b]	1.57		4.16[a]	1.60	1-2 years ago	3.64[a,b]	1.47		4.21[c]	1.54		Exceeds cutoff	4.18*	1.72
							2+ years ago	3.59[c]	1.65		3.31[a,b,c]	1.55				
				.71			9.43***			2.73*			.51			
Belonging	None	2.05	1.54		2.67[a]	1.54	Past month	1.94[a]	1.29		1.78	1.12		Under cutoff	1.68	1.21
	Thoughts	2.24	1.59		2.41	1.41	Past year	1.93[b]	1.20		1.82	1.21				
	Multiple	1.89	1.24		1.77[a]	1.15	1-2 years ago	1.93[c]	1.16		1.76	1.24		Over cutoff	1.95*	1.36
							2+ years ago	1.44[a,b,c]	1.06		1.62	1.11				
				46.67***			32.95***			3.47*			2.29			
Stigma	None	3.13[a]	2.17		3.10[a,c]	1.53	Past month	5.57[a]	1.27		5.43	1.37		Under cutoff	4.92	1.71
	Thoughts	3.45[b]	1.74		4.27[b,c]	1.67	Past year	5.18[a]	1.39		5.27	1.36				
	Multiple	5.42[a,b]	1.34		5.33[a,b]	1.41	1-2 years ago	5.27	1.42		5.38	1.61		Over cutoff	5.36*	1.54
							2+ years ago	5.38	1.41		4.95	1.59				
				.38			6.60***			3.57*			3.09*			
Interpersonal influence	None	2.23	1.49		2.97[a]	1.52	Past month	2.40[a]	1.37		2.09	1.15		Under cutoff	1.97	1.25
	Thoughts	2.60	1.56		2.71	1.37	Past year	2.36[b]	1.32		2.33	1.27				
	Multiple	2.31	1.31		2.16[a]	1.23	1-2 years ago	2.15	1.10		2.41	1.58		Over cutoff	2.39***	1.43
							2+ years ago	1.83[a,b]	1.04		1.91	1.17				

Note. BPD = Borderline Personality Disorder. For each subscale, values that share superscripts within each column statistically significantly differ from one-another ($p < .05$, 1-tailed). * = $p < .05$ (1-tailed); *** = $p < .001$ (1-tailed); all post-hoc analyses applied the Games-Howell post-hoc test, which does not assume equal group sizes or homogeneous variances. Sample sizes for History of Self-Injurious Thoughts and Behavior: Sample 5 SABS None = 15, Thoughts = 222, Attempt = 84, Multiple = 343; NSIBS None = 26, Thoughts = 14, Multiple = 624; Sample 6 SABS None = 18, Thoughts = 222, Attempt = 81, Multiple = 329; NSIBS None = 24, Thoughts = 22, Multiple = 604. Sample sizes for Recency of Self-Injurious Behavior: Sample 5 SABS Past month = 65, Past year = 124, 1-2 years ago = 58, 2+ years ago = 180 (total = 427); NSIBS Past month = 340, Past year = 162, 1-2 years ago = 65, 2+ years ago = 57 (total = 624); Sample 6 SABS Past month = 44, Past year = 123, 1-2 years

ago = 89, 2+ years ago = 154 (total = 410); NSIBS Past month = 320, Past year = 171, 1-2 years ago = 45, 2+ years ago = 68 (total = 604); Sample sizes for Probable BPD diagnosis: 116 under cutoff, 548 over cutoff.

5.4.6. Incremental Validity

The incremental validity of the SABS was explored in several hierarchical regressions involving a range of variables. Given that the subscales of the SABS are significantly correlated and, therefore, not completely independent, these analyses allowed us to identify which subscales showed unique, incremental power. Statistically significant univariate predictors were included in our multivariate models[11].

Depression, Anxiety Problems, and Optimism. We tested the incremental validity of the SABS in predicting lifetime SA thoughts and behavior above and beyond depressive symptoms (PHQ-8, IDAS-Dysphoria), anxiety problems (GAD-7), and optimism (LOT-R). Fully standardized coefficients are reported for logistic models (β_{stdXY}) to enable the relative magnitude of effects to be examined (Menard, 2011). These analyses (see Supplementary material for Chapter 5) demonstrated that the *Stigma* (β_{stdXY} = .96, p < .05, OR = 2.60) and *Revenge* (β_{stdXY} = -.81, p < .05, OR = .45) subscales statistically significantly predicted lifetime suicidal thoughts, and that the SABS *Belonging* (β_{stdXY} = -.39, p < .05, OR = .68), *Stigma* (β_{stdXY} = .37, p < .05, OR = 1.44, *Self-punishment* (β_{stdXY} = .63, p < .001, OR = 1.88), and *Revenge* (β_{stdXY} = -.52, p < .05, OR = .59) subscales statistically significantly predicted lifetime suicidal behavior. When controlling for the NSIBS, the *Stigma* subscale statistically significantly predicted lifetime suicidal thoughts (β_{stdXY} = 1.36, p < .05, OR = 3.88), and the *Belonging* (β_{stdXY} = -.66, p < .05, OR = .52), *Self-punishment* (β_{stdXY} = .88, p < .001, OR = 2.40), and *Revenge* (β_{stdXY} = -.53, p < .05, OR = .59) subscales statistically significantly predicted lifetime suicidal behavior.

It was not possible to test the incremental validity of the NSIBS in similar analyses because the PHQ-8, IDAS-Dysphoria, GAD-7, and LOT-R were all non-significant univariate predictors and were therefore excluded from our multivariate models. These results do usefully demonstrate, though, that depressive symptoms, anxiety problems, and optimism are relatively more specific to SA than NSSI.

Current Suicidal Thinking. We also explored whether the SABS provided incremental validity in predicting current suicidal thinking over and above a range of existing, theoretically important clinical constructs, and markers of well-being (Table 10). These analyses were conducted using two different measures of current suicidal thinking in

[11] Because there is some debate as to whether it is useful to exclude nonsignificant univariate predictors from multivariate models, the multivariate regressions were rerun using the excluded variables. The nonsignificant predictors did not demonstrate statistically significant unique associations, change the significance of variables already in the model, or meaningfully increase R^2.

Samples 5 and 6. Although we did not have theoretical reasons for examining demographic variables as predictors, we considered these for inclusion in the multivariate regressions in order to further clarify substantive relationships between demographic variables and our newly developed scales. Demographic variables were entered at Step 1, clinical and well-being variables at Step 2, and the SABS subscales were entered at Step 3.

The first of our linear regressions revealed that the SABS demonstrated a statistically significant ΔR^2 of .04 in predicting total scores on the BSS when controlling for demographic variables, clinical variables (perceptions of unlovability, unbearability, burdensomeness, thwarted belongingness, symptoms of BPD, experiential avoidance, perceived stress), and social support. The *Belonging* (β = -.71, $p < .05$), *Self-punishment* (β = .84, $p < .05$), and *Dependence* (β = 1.44, $p < .001$) subscales statistically significantly predicted current suicidal thoughts.

In the second linear regression, the SABS demonstrated a statistically significant ΔR^2 of .07 in predicting total scores on the DSI-SS when controlling for demographic variables, clinical variables (difficulties in regulating negative and positive emotions, reasons for living, emotional reactivity, current mood), and well-being variables (satisfaction with life, subjective happiness, subjective vitality). The *Stigma* (β = .36, $p < .05$), *Escape* (β = -.35, $p < .05$), and *Dependence* (β = .92, $p < .001$) subscales statistically significantly predicted current suicidal thoughts.

Taken together, these analyses reveal that the SABS demonstrates incremental validity in predicting lifetime suicidal thoughts and behavior and current suicidal thoughts above and beyond a wide range of demographic, clinical, and well-being variables. Our linear regressions provided a robust test of incremental validity because they employed a continuous variable (which will have improved power; cf. MacCallum, Zhang, Preacher & Rucker, 2002) and included several theoretically important clinical covariates (e.g., perceptions of unlovability, unbearability, burdensomeness, thwarted belongingness, symptoms of BPD, experiential avoidance, difficulties in regulating negative and positive emotions, reasons for living, emotional reactivity). Our analyses demonstrate that the SABS could explain an additional 4% and 7% of variance in current suicidal thoughts when controlling for a wide range of demographic, clinical, and well-being variables. These figures exceed what might be hoped for in typical tests of incremental validity (Hunsley & Meyer, 2003).

Table 10 Hierarchical Multivariate Linear Regressions Comparing the Incremental Validity of the Suicide Attempt Beliefs Scale (SABS) and Nonsuicidal Self-Injury Beliefs Scale (NSIBS) Against Demographic Variables, Established Clinical Predictors, and Well-Being Variables

	Step 1				Step 2				Step 3			
Variables	*B*	*SE B*	*β*	*ΔR²*	*B*	*SE B*	*β*	*ΔR²*	*B*	*SE B*	*β*	*ΔR²*
	Beck Scale for Suicide Ideation Total Score (Sample 5)											
				.04***				.44***				.04***
Constant	37.39***	.98	.01									
Level of education	-.47	.33	-.07									
Income	-.88**	.30	-.13									
Marital status	-.81	.43	-.08									
Constant					28.09***	3.14	.00					
Level of education					-.02	.26	-.00					
Income					-.19	.24	-.03					
Marital status					-.13	.34	-.01					
SCS: Unlovability					.19***	.04	.32					
SCS: Unbearability					.10*	.05	.11					
INQ: Burdensomeness					.22***	.04	.27					
INQ: Belongingness					-.05	.04	-.06					
BPD symptoms					-.09	.15	-.02					
Experiential avoidance					-.01	.03	-.03					
Perceived social support					-.10***	.02	-.19					
Constant									27.19***	3.06	-.00	
Level of education									-.07	.26	-.01	
Income									-.25	.23	-.04	
Marital status									-.22	.33	-.02	
SCS: Unlovability									.15***	.04	.25	
SCS: Unbearability									.08	.05	.09	
INQ: Burdensomeness									.18***	.04	.22	
INQ: Belongingness									.04	.04	-.05	
BPD symptoms									-.06	.15	-.01	
Experiential avoidance									-.02	.03	-.05	

Perceived social support									-.08***	.02	-.16
SABS: Belonging									-.61*	.31	-.09
SABS: Stigma									.21	.21	.04
SABS: Self-punishment									.42*	.20	.10
SABS: Eliciting help									-.29	.21	-.06
SABS: Escape									.08	.21	.02
SABS: Dependence									.97***	.26	.17
SABS: Revenge									-.24	.26	-.04
	Depressive Symptom Inventory-Suicidality Subscale Total Score (Sample 6)										
			.05***				.27***				.07***
Constant	8.71***	.57	-.01								
Age	-.01	.02	-.05								
Employment status	.17	.12	.08								
Education level	-.34**	.13	-.13								
Income	.01	.13	-.00								
Marital status	-.29*	.15	-.09								
Constant				8.84***	1.33	-.01					
Age				-.01	.01	-.05					
Employment status				.10	.11	.05					
Education level				-.10	.12	-.04					
Income				.08	.11	.03					
Marital status				-.19	.13	-.06					
DERS: Nonacceptance				-.01	.02	-.03					
DERS: Goals				-.03	.03	-.04					
DERS: Impulse				.06*	.03	.14					
DERS: Awareness				-.02	.02	-.03					
DERS: Strategies				.04	.03	.09					
DERS: Clarity				.03	.03	.05					
DERS-P: Nonacceptance				.06	.03	.09					
DERS-P: Goals				.00	.03	.00					
DERS-P: Impulse				.00	.03	.01					
BRFLS: Responsibility				-.06	.04	-.07					

BRFLS: Children	-.03	.03	-.04			
BRFLS: Fear	-.01	.04	-.01			
BRFLS: Morality	-.21***	.05	-.19			
Emotion reactivity	.00	.01	-.01			
Current mood	.29	.17	.08			
Satisfaction with life	-.04*	.02	-.10			
Subjective happiness	-.01	.00	-.08			
Subjective vitality	-.03	.02	-.08			
Constant				7.53***	1.30	.00
Age				-.02	.01	-.07
Employment status				.06	.10	.03
Education level				-.04	.11	-.01
Income				.05	.11	.02
Marital status				-.22	.13	-.07
DERS: Nonacceptance				-.01	.02	-.02
DERS: Goals				.00	.03	.01
DERS: Impulse				.03	.03	.07
DERS: Awareness				-.02	.02	-.04
DERS: Strategies				.01	.03	.03
DERS: Clarity				.02	.03	.04
DERS-P: Nonacceptance				.04	.03	.06
DERS-P: Goals				-.02	.03	-.04
DERS-P: Impulse				.01	.03	.02
BRFL: Responsibility				-.05	.04	-.06
BRFL: Children				-.03	.03	-.05
BRFL: Fear				.01	.04	.01
BRFL: Morality				-.17*	.05	-.15
Emotion reactivity				.00	.01	.01
Current mood				.32*	.16	.09
Satisfaction with life				-.03	.02	-.08
Subjective happiness				-.01	.00	-.08
Subjective vitality				-.03	.02	-.08

SABS: Belonging	-.16	.13	-.06
SABS: Stigma	.21*	.09	.12
SABS: Self-punishment	-.02	.08	-.01
SABS: Eliciting help	.03	.09	.02
SABS: Escape	-.20*	.09	-.12
SABS: Dependence	.63***	.11	.31
SABS: Revenge	-.08	.11	-.04

Note. SCS = Suicide Cognitions Scale; INQ = Interpersonal Needs Questionnaire; BPD symptoms = Mclean Screening Instrument for Borderline Personality Disorder; Experiential Avoidance = Brief Experiential Avoidance Questionnaire; Perceived social support = Multidimensional Scale of Perceived Social Support; DERS = Difficulties in Emotion Regulation Scale; DERS-P = Difficulties in Emotion Regulation Scale-Positive; BRFLS = Brief Reasons for Living Scale; Emotion reactivity = Emotion Reactivity Scale; Satisfaction with life = Satisfaction with Life Scale; Subjective happiness = Subjective Happiness Scale; Subjective vitality = Subjective Vitality Scale.
The BRFL Survival subscale was excluded from the regression. * = $p < .05$ (2-tailed); ** = $p < .01$ (2-tailed); *** = $p < .001$ (2-tailed)

5.5. Discussion

The SABS and NSIBS represent a significant advance in that they display very promising psychometric properties and appear to measure content beyond what is contained within existing instruments. These analyses also further demonstrate that the SABS and NSIBS need to be understood as related but separate measures: that beliefs about SA and NSSI are similar but separate. Both instruments demonstrated "moderate" to "excellent" test retest reliability over 2-4 weeks, consistent with our aim to measure beliefs (relatively enduring personal meanings for SA and NSSI that confer vulnerability across situations but which also potentially fluctuate or change due to a range of factors; Beck et al., 1979).

Both instruments also demonstrated strong internal consistency. The AICs provided particularly useful information. They indicated that the subscales of the SABS and NSIBS measure relatively specific domains of SICs and that each instrument as a whole taps a range of distinct, moderately related content. We were particularly interested to find that many of the SABS subscales demonstrated virtually no relationship with the NSIBS subscales, and that even subscales that have identical labels across both instruments only evidence "moderate" to "large" correlations (Cohen, 1998). These results corroborate the structural analyses presented in Part I and further evidence that the SABS and NSIBS each measure separate constructs and are best understood as separate measures.

We also presented a large range of correlations to support the convergent and divergent validity of the SABS and NSIBS. Three general findings emerged. First, the SABS and NSIBS each demonstrated "small" to "moderate" (Cohen, 1998) positive correlations with a range of clinical variables (depressive symptoms, anxiety symptoms, symptoms of BPD, experiential avoidance, perceived stress, difficulties in regulating positive and negative emotions, emotional reactivity) and "small" to "moderate" (Cohen, 1998) negative correlations with several measures of well-being (optimism, perceived social support, satisfaction with life, and subjective happiness and vitality). Second, correlations between the subscales of the SABS and NSIBS and other variables were sometimes similar in magnitude and sometimes differed by up to $r = \sim.15$. This evidence further supports the notion that the SABS and NSIBS measure similar but separate content.

Third, an important step in the development of any new instrument involves exploring associations between the newly developed instrument and existing, similar measures. The SABS and NSIBS both demonstrated "small" to "moderate" (Cohen, 1998) correlations with purportedly similar, existing constructs, including the reasons for SA and NSSI, reasons to live, current suicidal thinking, and perceptions of unlovability, unbearability,

burdensomeness, and thwarted belongingness. Critically, these results indicate that beliefs about SA and NSSI appear to be different to these existing constructs, at least in terms of the measures we used here. In contrast, the IMSA (May & Klonsky, 2013), for example, demonstrated large correlations with other measures of the reasons for SA, some of which were sufficiently large to indicate equivalence (<.75; Clark & Watson, 1995).

We also found that the SABS significantly incrementally predicts previous and current self-injurious cognitions and behavior over and above demographic characteristics and the clinical and well-being variables just discussed. Taken together, these findings further evidence that our scales relate in expected ways to a range of variables and suggest that the SABS and NSIBS represent a significant advance in that each measures content that is not currently within the measurement spotlight.

Finally, we found that participants' increasingly endorsed the SABS and NSIBS the more they had progressed from thoughts to taking action to engaging in multiple episodes of SA or NSSI, and the more recently they had made a SA or engaged in NSSI. These findings are consistent with cognitive theories that link the strength of SA cognitions to greater risk of SA (e.g., Rudd, 2000, 2006). They indicate that the SABS and NSIBS may be sensitive to changes over time and potentially implicate beliefs about SA and NSSI in the development and repetition of SA and NSSI, although a direct test of this hypothesis is of course required before any firm conclusions can be drawn.

Limitations and Future Directions

SA and NSSI beliefs convey what SA and NSSI *means* to each individual. They contrast with variables such as gender, age, or life events, which cannot be modified. In Part I, we proposed that strong endorsement of SICs may be a specific, defining, and time-varying feature of SA and NSSI, and may potentially explain why some people, at some times, engage in these behaviors specifically, rather than choosing to use an alternative self-regulatory strategy (e.g., listen to music; go for a walk; problem-solve; ruminate; get drunk) in response to a particular internal or external trigger (e.g., emotional distress; social exclusion). This hypothesis is grounded in Beck's (1976) concept of cognitive content specificity: the idea that the presence of *particular* cognitions differentially and uniquely predicts the development and maintenance of specific psychological problems. For example, the cognitions that characterise depression have been found to be different from those that characterise posttraumatic stress disorder (Ehring, Ehlers, & Glucksman, 2008). We see SA and NSSI beliefs as being a central component of a relatively transient mode of mind that becomes strongly activated for short periods of time. This state of mind probably involves

corresponding information-processing biases and links between cognition, emotion, and behavior (cf. Beck & Haigh, 2014).

From a cognitive-behavioral perspective (Beck, 2011; Powers, 2005; Young, Klosko & Weishaar, 2003), SIB is viewed as a response to the presence and activation of SA and NSSI beliefs. It follows that (i) whether or not a (relatively transient) intention to self-injure is formed; (ii) the specific, idiosyncratic details of each SIB episode (e.g., trigger, method, when to stop, function, onset, repetition); and (iii) whether each episode involves suicidal or nonsuicidal intent, would all be expected to be linked to or driven by activation of idiosyncratic SA and NSSI beliefs – which the SABS and NSIBS measure.

Whilst we believe that these hypotheses seem plausible, they are untested. The starting point for investigating such questions was developing the SABS and NSIBS; further research is now needed to test whether, under what circumstances, and for whom, SA and NSSI beliefs cause or drive (future) SIB, and to pit the predictive ability of the SABS and NSIBS against other measures (e.g., defeat/entrapment; Siddaway, Taylor, Wood & Schulz, 2015). Establishing the incidence of SA and NSSI beliefs (rather than the presence/absence of suicidal thoughts or behavior) would extend current epidemiological research.

We were surprised that the SABS and NSIBS generally evidenced "small" negative correlations (Cohen, 1998) with the BRFLS. It is possible that these results occurred because the BRFLS has just two items per subscale, or because the items of the BRFLS appear to measure reasons to live (e.g., "I believe I can find a purpose in life, a reason to live") and reasons for not making a SA (e.g., "I do not want to die"), which may have clouded relationships. Correlations between the SABS and NSIBS and reasons to live need replicating using the full-length Reasons for Living Inventory (Linehan, Goodstein, Nielsen, & Chiles, 1983), although this scale also appears to contain items that tap a mixture of reasons to live and reasons not to die (which may be different to one-another). We presented a large number of correlations without any correcting for chance and it is therefore possible that some of our significance results may not replicate.

It is important to know what predicts the deactivation and replacement of SA and NSSI beliefs. We found that being in a deactivated positive mood (feeling peaceful/serene, relaxed/calm) was associated with a lower risk of endorsing some subscales of the SABS and NSIBS. This result provides tentative evidence for the potentially protective role of calmness in deactivating SICs (see Siddaway, Taylor & Wood, in press) but further research, particularly ecological momentary assessment, is now needed to explicate the presumably dynamic relationship between mood changes and the activation and deactivation of SICs.

Further research on the incremental validity of our scales is urgently needed, particularly for the NSIBS as its incremental validity is currently untested. Such analyses would preferably use continuous variables to improve power. The specificity of the SABS and NSIBS might be examined by administering these instruments to theoretically-distinct groups. For example, it would be particularly interesting to compare endorsement of the SABS and NSIBS in individuals who have thought about or engaged in SA and/or NSSI with individuals who have thought about or engaged in (i) behaviors which immediately but unintentionally result in physical injury (e.g., drink driving; skin picking; nail biting; hair-pulling), and (ii) behaviors in which physical injury may sometimes be a delayed or unanticipated side-effect (e.g., smoking, over-eating, binge drinking, or unprotected sex).

CHAPTER 6

6. General Discussion

6.1. Overview

This thesis consists of four core chapters. Each chapter is a manuscript that has either been published (Chapters 2 and 3) or submitted for publication (Chapters 4 and 5) to a peer-reviewed psychology journal. This thesis sought to test the two central hypotheses of PCP, that most psychological constructs are bipolar, consisting of a positive pole and a negative pole, and that positive constructs provide important unshared explanatory power in relation to psychological problems (Johnson & Wood, 2017; Joseph & Wood, 2010; Wood & Johnson, 2016; Wood & Tarrier, 2010). It also aimed to make a contribution in its own right by clarifying PCP's core messages and making them more explicit, as PCP has not always been explicit or consistent in stating and discussing its core predictions (Held, 2016).

6.2.Summary of Thesis Results and Their Implications for Positive Clinical Psychology

6.2.1. Chapters 2 and 3

Chapters 2 and 3 examined whether two popular measures of depressive symptoms and anxiety problems – the CES-D and STAI – could be reconceptualised, respectively, as measures of a depression-well-being continuum and an anxiety-calmness continuum (Joseph & Wood, 2010; Tarrier & Wood, 2010). The CES-D and STAI each contain a mixture of negatively worded items (which, respectively, measure the presence of depressive and anxiety symptoms) and positively worded items (which, respectively, measure the presence of well-being and calmness). It was this key feature of each scale that allowed them to be re-examined as bipolar continua.

Chapters 2 and 3 provide evidence that depression needs to be understood as one pole of a depression-well-being continuum and anxiety problems need to be understood as one pole of an anxiety-calmness continuum. These findings have important implications for the structure and definition of depression and anxiety, as they go against the view that depressive and anxiety symptoms range from zero to intense.

In this respect, the CES-D and STAI offer measurement properties beyond scales that do not contain a mixture of positively and negatively worded items. For example, scores on the Beck Depression Inventory (BDI-II; Beck et al., 1996) have a potential range of 0 to 63, with higher scores indicating greater depression. This scale does not contain any positively worded items that need to be reverse-scored, as is customary in many scales. A score of zero

on the BDI-II therefore indicates the absence of depressive experiences, but it does not indicate the presence of the opposite of depression. However, for two people scoring zero on the BDI-II, one could be high on well-being, the other low (Joseph & Wood, 2010). Scales that measure depression or well-being, or anxiety or calmness, but which do not contain a mixture of positively and negatively worded items, are presumably measuring half of the depression-well-being continuum or the anxiety-calmness continuum.

PCP has suggested that most psychological constructs are bipolar, consisting of a positive pole and a negative pole. However, there is little direct evidence to support this claim and the authors placed a large amount of importance (Joseph & Wood, 2010; Wood & Tarrier, 2010) on two studies they conducted (Joseph & Lewis, 1998; Wood, Taylor & Joseph, 2010), which demonstrated that depression can be understood as forming a continuum with well-being. Joseph and Wood (2010) also made a theoretical argument that anxiety forms a continuum with calmness in order to support the notion that most psychological constructs are bipolar. However, no direct evidence was provided to support this claim. Chapters 2 and 3 therefore provide much-needed, important evidence regarding whether bipolar continua exist and are consistent with the hypothesis that most psychological constructs are bipolar.

Chapters 2 and 3 also separately explored the (linear or nonlinear) form of the relationship between the depression-well-being continuum and other psychiatric variables and the anxiety-calmness continuum and other psychiatric variables for the first time. Baseline levels on each continuum were (separately) found to have a near linear relationship with psychiatric outcome variables measured at the same time and one and two years later. These results indicate that there is no intrinsic way to demarcate problematic degrees of depressive or anxiety symptoms (or beneficial degrees of well-being or calmness) based on how the facets of each continuum are related to other psychiatric variables. That these results were apparent over time suggests that moving along each continuum toward high well-being and high calmness provides continuous and long-term protection against other psychological problems.

These results point to the joint importance and usefulness of treating depressive and anxiety symptoms and promoting well-being and calmness. As discussed in Chapters 2 and 3, these results also point to the usefulness of early intervention and prevention (when people begin to move away from high well-being and high calmness) and instilling resilience (by providing interventions to move people toward high well-being and high calmness). Fostering high levels of well-being and calmness would mean that individuals have further to go before

they reach high levels of depressive and anxiety symptoms. Psychiatric and psychological interventions that are grounded in a continuum conceptualization would logically be stopped when an individual reports high well-being or calmness. However, it must be explicitly noted that none of these predictions were tested directly. For example, in a clinical trial, or by using ecological momentary assessment to monitor changes in well-being or calmness over time and how those changes relate to changes in other psychological problems. Chapters 2 and 3 call for a direct test of these questions.

In exploring the (linear or nonlinear) form of the relationship between the depression-well-being continuum and other psychiatric variables and the anxiety-calmness continuum and other psychiatric variables, Chapters 2 and 3 provide much needed evidence to extend and test the very minimal discussion provided by PCP regarding the (linear or nonlinear) form of bipolar continua (Joseph & Wood, 2010; Wood & Tarrier, 2010). The findings presented in these chapters raise the possibility that all bipolar continua will demonstrate near linear relationships with other variables. This possibility needs empirically testing.

6.2.2. Chapters 4 and 5

Chapters 4 and 5 reported the development and validation of the SABS and NSIBS, two new self-report measures of beliefs about SA and NSSI. The evidence presented in these chapters support the use of separate terminology and definitions for SA and NSSI, and indicate that SA and NSSI need to be separated in research designs. This is important because for over 50 years, many researchers, particularly those in the UK, have been treating SA and NSSI as one and the same construct.

Item generation was guided by the broad hypothesis that SICs are characterised by positive and negative SICs. “Positive SICs” were hypothesised to be cognitions about the individual and interpersonal perceived advantages of SA and NSSI (e.g., “Attempting suicide changes the way that I am thinking” [individual]; “NSSI helps me fit in with other people” [interpersonal]). “Negative SICs” were hypothesised to be cognitions about the individual and interpersonal perceived disadvantages of SIB (e.g., “NSSI makes my problems worse” [individual]; “People think that my suicide attempt(s) are selfish” [interpersonal]).

As described in Chapter 4, the majority of the subscales of the SABS and NSIBS appear to measure positive SICs (cognitions about the perceived advantages of SIB) and few of the subscales appear to measure negative SICs (cognitions about the perceived disadvantages of SIB). CFA analyses revealed that the SABS and NSIBS are each most appropriately understood as a series of correlated but separate subscales/factors (separate facets of each individual’s mental representation of SIB).

Although the prediction that SICs are characterised by positive and negative SICs seemed relatively straightforward, in practice, the process of scale development revealed the complexities involved in measuring and understanding SICs. For this reason, as described in Chapter 4, the author deliberately refrained from claiming that particular SABS or NSIBS subscales measure either positive or negative SICs. It seemed conceivable, for instance, that some respondents may perceive that a particular SABS or NSIBS item describes an advantage of SIB for them whilst other respondents describe the same item as conferring a disadvantage of SIB for them. It also seemed possible that some subscales contain positive and negative SICs. As discussed in Chapter 4, further research is needed to make sense of the function of particular items, and to explore whether people interpret certain items differently. This might be achieved, for example, by asking each individual whether specific items represent an advantage or a disadvantage of SA or NSSI for them, or by testing how the activation of particular items increases or decreases risk for SA or NSSI by linking activation of particular SABS and NSIBS items to changes in the strength of SA and/or NSSI urges, desire, or intent.

Chapter 5 provides some evidence in relation to PCP's second prediction, that "positive" constructs provide important unshared explanatory power in relation to psychological problems (Johnson & Wood, 2017; Joseph & Wood, 2010; Wood & Johnson, 2016; Wood & Tarrier, 2010). During the process of validating the SABS and NSIBS, several measures of well-being were collected (optimism, perceived social support, satisfaction with life, and subjective happiness and vitality). These were included in tests of the incremental validity of the SABS, predicting current suicidal thinking. In (separate) multivariate analyses, optimism, perceived social support, and satisfaction with life all demonstrated statistically significant, unique predictive relationships with current suicidal thinking when controlling for a range of theoretically relevant, negative constructs. These results are consistent with PCP's claim that variables that particularly characterise well-being, and which have been particularly studied by positive psychology, will provide incremental validity in predicting psychological problems. PCP recognises that this would be the case because these variables are different to variables studied by mental health researchers and therefore share less variance (Johnson & Wood, 2017).

Chapters 4 and 5 did not directly test whether SA and NSSI cognitions have an opposite pole, being part of a bipolar continuum (Johnson & Wood, 2017; Joseph & Wood, 2010; Wood & Johnson, 2016; Wood & Tarrier, 2010). The reason for this is that there does not seem to be a construct that might theoretically be the opposite of SA and NSSI

cognitions. Put differently, SA and NSSI cognitions seem to be unipolar rather than bipolar. The *absence* of SA and NSSI cognitions is measured by the SABS and NSIBS (respondent's select "Strongly disagree" to items and obtain a low score). Rephrasing SABS and NSIBS items (e.g., "I attempt suicide because I deserve to suffer" "NSSI helps me escape negative emotions") would not measure the *presence* of the *opposite* of cognitions about SA and NSSI. For example, the item "I attempt suicide because I deserve to suffer" could be rephrased to become "I attempt suicide because I *do not* deserve to suffer" and "I *do not* attempt suicide because I deserve to suffer." These two new items convey entirely different, nonsensical meanings, and neither item would create a bipolar measure of deserving to suffer in relation to SA, for example.

The SABS and NSIBS generally evidenced "small" negative correlations (Cohen, 1998) with reasons for living, which goes against the idea that the SABS and NSIBS measure the opposite of reasons for living (very large, negative correlations would be expected). The SABS and NSIBS measure beliefs about SA and NSSI, which are related to but separate but from reasons for living. The SABS and NSIBS were not hypothesised to be opposites to reasons for living, although this possibility might arise in some reader's minds.

6.3. Are Positive and Negative Constructs Different Facets of the Same Variables?

The evidence presented in this thesis is consistent with PCP's prediction that most psychological constructs are bipolar, consisting of a positive pole and a negative pole. This thesis demonstrated that some constructs have bipolar opposite whilst other constructs do not.

Chapters 2 and 3 demonstrated that variables that particularly characterise well-being ("positive constructs;" in this case well-being and calmness) can form a continuum with variables that particularly characterise psychological problems ("negative constructs;" in this case depressive and anxiety symptoms). Depressive and anxiety symptoms have fairly intuitive and logical theoretical opposites. To illustrate, most diagnostic criteria for depression involve high negative affect, low positive affect (anhedonia), lack of engagement, poor sleep, impaired appetite, and poor social relationships (Wood & Johnson, 2016). Each of these symptoms is clearly on continua that range from positive to negative, respectively low negative affect, high positive affect, engagement, good sleep, appropriate eating, and good social relationships (Wood & Johnson, 2016).

Chapters 4 and 5 demonstrate that variables that particularly characterise psychological problems ("negative constructs;" SA and NSSI cognitions) are sometimes unipolar. There is no obvious opposite to SA and NSSI cognitions. Whether or not bipolar

continua characterise *most* psychological constructs, as PCP predicts, remains an empirical question.

6.3.1. Accounting for Complexity

These mixed findings illustrate the complexities involved in integrating variables that particularly characterise well-being ("positive constructs") with variables that particularly characterise psychological problems ("negative constructs") to explain psychological problems. The different studies reported here illustrate the fact that any test of PCP's hypothesis of bipolarity will be entwined with semantics (phrasing of items), the scale used (e.g., whether a scale contains a mixture of positively-worded and negatively-worded items that load on separate factors), theory and the construct in question (any empirical test of bipolarity must be underpinned by a theoretical prediction and rationale regarding how particular "positive" and "negative" constructs are opposite poles of the same variable), and measurement considerations (whether a particular bipolar construct can be measured). Including positively and negatively worded items in a single scale allowed plausible positive-negative continua for depression and anxiety problems to be tested using CFA in Chapters 2 and 3, but this is only one means of examining bipolarity and a consensus picture is needed to firmly establish evidence of bipolarity for particular constructs.

6.3.2. A Revision to Positive Clinical Psychology's Central Hypothesis?

The different findings presented in this thesis lead to the following conclusion: *Some psychological constructs are bipolar whilst other constructs are unipolar*. The following two predictions appear to explain the varied findings presented, and extend PCP's central prediction:

(1). Psychological constructs that are common in the general population – such as depressive symptoms, anxiety symptoms, well-being symptoms, and calmness symptoms – appear to be bipolar, having a positive and a negative pole;

(2). Psychological constructs that are rare in the general population and which specifically characterise psychological problems (rather than being an extreme manifestation of a common psychological experience) – such as SA and NSSI cognitions – appear to be unipolar.

The second prediction might be broadened by not specifying unipolarity only in relation to certain psychological problems and instead stipulating that *psychological constructs that are rare in the general population appear to be unipolar*. For example, Maslow (1943) predicted that "self-actualisation" is relatively rare in the general population and is the final level of psychological development that can be achieved when all basic and

psychological needs are fulfilled and the "actualization" of an individual's full personal potential has occurred. However, this thesis is concerned with psychological problems ("negative constructs" or variables that particularly characterise psychological problems), hence the phrasing of prediction two.

6.4. Future Directions for Positive Psychology and Positive Clinical Psychology

PCP and positive psychology are both relatively new literatures. Both make predictions regarding the relationship between variables that particularly characterise well-being ("positive" constructs) and those that particularly characterise psychological problems ("negative" constructs). Positive psychology attempted to carve out a niche through separatism, arguing that the constructs it examined ("positive" constructs) had been generally neglected by existing research and indeed are distinct from "negative" constructs. PCP also attempted to create its own niche but through calling for positive psychology research to be integrated with research on mental health problems. PCP argued that "positive" and "negative" constructs are often the same variable, viewed through different lenses. The relationship between "positive" and "negative" forms the core of both literatures and has implications for measurement, research designs, and interventions. It is therefore surprising that such a central issue has received relatively little research attention.

It is an empirical question whether, how, for whom, and under what circumstances, variables that particularly characterise well-being ("positive constructs") do or do not form bipolar continua with variables that particularly characterise psychological problems ("negative constructs). Several authors have asserted that "positive constructs" are separate from "negative constructs" (e.g., Duckworth et al., 2005; Wong, 2011) whereas PCP asserts that these constructs form bipolar continua. This thesis looked at a few of these questions and the findings suggest that some constructs are bipolar and others are unipolar.

Of course, when positive and negative constructs are separate, correlated facets of a single higher-order construct, this raises the possibility that any associations found between each construct and other variables may be attributable to the opposing (positive or negative) construct. In this context, interaction effects need to be considered between positive and negative constructs because, statistically, main effects cannot be interpreted in the presence of an interaction (Wood & Johnson, 2016). This was found to be the case by Kelly et al. (2011) in relation Bipolar Disorder, for example, where positive and negative appraisals of the same mood states interact such that the effects of positive appraisals are very different when negative appraisals are also considered.

One possibility may be that all constructs are bipolar at a particular level of abstraction, as Joseph and Wood (2010) discuss in some depth. For example, as discussed above, SA and NSSI cognitions do not have an obvious opposite construct. However, the constructs predicted to *underlie* SA cognitions do have intuitive opposite constructs. For example, hopelessness is strongly linked with the development of SA cognitions by several theories of SA (e.g., Rudd, 2000; Rudd et al., in preparation; Wenzel et al., 2009) and has an obvious opposite – hope*ful*ness – hope. A similar construct to hope, optimism, is already understood as being a bipolar optimism/pessimism construct (Carver et al., 2010), as measured by the Life-Orientation Test Revised (Scheier et al., 1994). As Carver et al. (2010) note "People range from very optimistic to very pessimistic, with most being somewhere between" (p. 880).

As the replication crisis literature illustrates, any one finding using one research design cannot definitively support a particular conclusion. Much more research is therefore needed in order to create a consensus of evidence on the relationship between positive and negative constructs, and to clarify which constructs are bipolar. The depression-well-being and anxiety-calmness continua need to be demonstrated using other measures and research designs to those used here, as do other hypothesised bipolar continua.

Joseph and Wood (2010) highlight how it may be possible to reinterpret existing measures of well-being or psychological problems if those measures happen to contain a mixture of positively and negatively worded items and positively and negatively worded items load on separate, correlated factors. However, it will probably be necessary to develop new measures in order to test whether most psychological constructs are bipolar. Since the personality and social psychology literature has long-debated the uni- versus bipolarity of affect (see Chapter 2), it will be important to consult that literature and transfer advances and key research methods to the literature investigating the potential bipolarity of psychological problems. Some simple but critical recommendations from that literature, for instance, are to take into account semantics when testing for bipolarity and to ensure that each pole of a continuum is measured by an equal numbers of items (Barrett & Russell, 1998). Bringing these factors together, one might develop and validate a scale that measures a particular hypothetical continuum using equal number of items that are semantic opposites and opposites in terms of valence (positive versus negative).

It might be the case that particular constructs are bipolar but that this fact has yet to be recognised in existing research because the complexity of human life and human psychological problems has clouded this issue. PCP argues that the bipolarity of positive and

negative constructs has implications for prediction and interventions (Johnson & Wood, 2017; Joseph & Wood, 2010; Wood & Johnson, 2016; Wood & Tarrier, 2010). These are useful suggestions. However, there are obvious practical barriers to overcome. For example, it is generally only practical to administer a set number of measures in a clinical trial and it is obviously most informative to measure what one theoretically intends to change. Most commonly, this will be the reduction of psychological problems, as measured by so-called "negative constructs." It may be the case that getting rid of psychological problems does lead to the presence of well-being, but only for particular people or problems, or under particular conditions, and these possibilities need testing.

The above discussion and this thesis has generally focused on the uni- versus bipolarity of "positive" and "negative" constructs. This discussion rests on the assumption that "positive" and "negative" constructs can be meaningfully distinguished. Held (2016, 2017) has provided thoughtful critiques of this assumption and its implications and raised issues that will need to be addressed if PCP and positive psychology are to continue to evolve. She points out that positive psychology and PCP have both made the a prior assumption that "positive" constructs are *adaptive* and *desirable* whereas "negative" constructs are *maladaptive* and *undesirable*; she points out that this assumption is simplistic, inaccurate, and, in fact, potentially undermines the entire foundation for positive psychology and PCP (Held, 2016, 2017). She goes on to provide an extended argument to illustrate that the language of "positive" and "negative" is arbitrary and dependent on one's perspective and context, particularly focusing on the problems associated with assuming that "negative" variables have no useful role to play (garnering attention only in order to be reduced or eliminated).

She states: "Whether any one emotion is positive or negative will…vary both within and between persons, depending on the goals in play in relation to a particular circumstance in any given moment. For example, if one seeks to end a relationship experienced as destructive, then anger may be consistent with that evaluation and thus aid in attaining that goal (and so may feel pleasant). Therefore, we cannot say whether any emotion is positive or negative prior to knowing a person's particular circumstances, especially the challenges that are faced in relation to needs, goals, and values in any given life situation" (Held, 2017, p. 10).

Another example might be guilt, an aversive emotion, but one which actually serves to motivate people to make up for their mistakes in order to repair meaningful bonds and maintain social relationships (Baumeister, Stillwell & Heatherton, 1994). Baumeister,

Stillwell and Heatherton (1995) present two studies which demonstrate that feeling guilty is associated with higher rates of learning lessons, changing subsequent behaviour, apologizing, confessing, and recognizing how a relationship partner's standards and expectations differ from one's own. Inducing guilt was found to be a costly but effective way of influencing the behavior of relationship partners. These two examples (and many others could be offered) illustrate that supposedly "negative" moods actually confer highly useful and functional value, depending on the context. In fact, a large literature supports the notion that emotions of any kind are not inherently good or bad; evidencing that positive and negative emotions are equally functional as they both provide important information that can be used to guide behaviour and respond effectively to the environment (see Baumeister et al., 2007). Furthermore, a book entitled *The Positive Side of Negative Emotions* presents three decades of research demonstrating how various "negative" emotions such as sadness, anxiety, jealousy, and guilt and "negative" coping styles such as defensive pessimism, can be functional/adaptive.

Held's (2016, 2017) concerns regarding whether "positive" and "negative" constructs can be meaningfully distinguished and what language is most useful in accounting for context seem valid concerns. However, it is not clear whether her concerns are stated too strongly as PCP does, at times, recognise the importance of considering that the functional value of "negative" characteristics. For example, Wood and Tarrier (2010) wrote that "any designation of a characteristic as positive or negative is simplistic and inaccurate, as any trait or emotion can be 'positive or negative' depending on the situation and concomitant goals and motivations" (p. 827). Chapters 2 and 3 of this thesis did not explicitly focus on the potential advantages that being anxious or depressed may offer. The adaptiveness of anxiety is well-known and intuitive as a survival mechanism.

The adaptiveness of depression is less immediately apparent but has been discussed by several authors. For example, according to social rank theories, depression arises as an evolutionarily adaptive, short-term threat-defence response in situations where an individual perceives that they are unable to escape from a state of being defeated (Sloman, 2000; Sloman, Gilbert & Hasey, 2003). The symptoms of depression (e.g., low positive affect, negative self-referent cognitions, behavioural inhibition) are thought to function as a damage limitation strategy in order to avoid pursuing goals that cannot be obtained or would decrease the ability to survive and reproduce if they were pursued regardless of the danger or cost (Gilbert, 1998). Social rank theories suggest that psychological problems can emerge from

the malfunction of the usually evolutionarily adaptive threat-defence response (e.g., Gilbert, 1998; Sloman et al., 2003).

Held (2016, 2017) and McNulty and Fincham (2012) both advocate abandoning the false dichotomy of "positive" and "negative" labelling of variables and assert that the function and adaptiveness of a construct depends on a complex array of factors (context). Held (2016) prefers the terms "constructive" and "destructive" (or "adaptive" and "maladaptive") because they hint at the necessity to consider each person in their context.

McNulty and Fincham (2012) highlight that "the psychological characteristics that benefit people experiencing optimal circumstances may not only fail to help people experiencing suboptimal circumstances, but may harm them" (p. 106). They demonstrate empirically how forgiveness, optimism, benevolent attributions, and kindness – favourite "positives" of positive psychologists – can diminish adaptive functioning.

Held (2016) discusses the importance of context in relation to "positive interventions" (discussed further below) and argues that if the distinction between positive and negative characteristics is arbitrary, simplistic, and inaccurate, then it follows that the distinction between "positive" and "negative" interventions is also arbitrary, simplistic, and inaccurate. "Positive" and "negative" interventions are cast as two distinct kinds of interventions by positive psychology and, at times, by PCP; however, this conceptualization may be more apparent than real (Held, 2016, 2017). For example, Wood, Perunovic, and Lee (2009) report evidence that repeating positive self-statements (a "positive" intervention) such as "I accept myself completely" or focusing on ways in which the statement was true caused those with low self-esteem to feel worse and boosted those with high self-esteem only mildly.

A final concern Held (2016, 2017) raises against PCP is that PCP does not clarify how positive and negative constructs may be properly "balanced" in the service of optimal functioning, even though they repeatedly call for a "balance" of "positive" and "negative" constructs. For positive psychology or PCP to continue to evolve, these potential concerns will need to be fully addressed.

6.5. "Positive Interventions"

PCP argues positive constructs can be fostered to treat psychological problems because these constructs are different to those already in common use in relation to psychological problems. Johnson and Wood (2017) and Wood and Johnson (2016) provide useful summarises of applying "positive constructs" to treat psychological problems. These have been termed "positive interventions" to denote the idea that these interventions have differed in their focus from existing mental health treatments. A few observations can be

made that extend the discussion presented by Johnson and Wood (2017) and Wood and Johnson (2016).

First, it may that "positive" constructs (variables that particularly characterise well-being and which especially been studied by positive psychology) can be used to alleviate "negative" constructs, either because the variables are different (e.g., gratitude and depression), or because interventions that increase the positive pole of a bipolar trait can be used to reduce the negative pole of that trait because *both poles are different facets of the same variable*. It is an empirical question whether particular "positive interventions" or mental health treatments are respectively more effective in promoting "positive" constructs or reducing "negative" constructs.

A burgeoning body of research suggests that "positive interventions" can enhance positive mood and well-being. For example, a meta-analysis by Bolier et al. (2013) examined whether "positive interventions" enhanced well-being and reduced depressive symptoms and found a small effect size for each. The finding that enhancing well-being reduces depressive symptoms seems relatively unsurprising, given that decades of research by David Watson and others have documented that depression is particularly characterised by low well-being (high negative affect and low positive affect; see Mineka, Watson & Clark, 1998; Watson, 2009). This meta-analysis indicates, however, that "positive interventions are *less* effective than existing, well-established treatments for depression. For example, a review of meta-analyses found that large effect sizes are evident in trials of CBT for depression (Butler, Chapman, Forman & Beck 2006).

As discussed in the Introduction, cognitive therapy for depression, for example, includes specific components such as behavioural activation that aim to bolster well-being and other components that aim to reduce depression (Beck et al., 1979). Cognitive therapy for anxiety problems (Beck, Emery & Greenberg, 1985) includes specific components that aim to reduce anxiety symptoms such as relaxation and find more realistic and helpful alternative cognitions to anxious predictions to promote confidence and self-efficacy. Cognitive therapy for low self-esteem includes components that specifically attempt to reduce self-criticism and components that specifically attempt to bolster engagement in meaningful relationships and mastery experiences as part of discovering and elaborating alternative, more compassionate ways to understand one's self and the world (i.e. positive constructs) (Fennell, 1997).

These examples (countless others are available) illustrate that existing, well-established (see Butler et al., 2006) interventions for psychological problems often already jointly aim to reduce negative constructs *and* increase positive constructs. Chapters 2 and 3

indicated that total CES-D and STAI scores provide an indication of depressive and anxiety severity that is based on a combination of the presence/absence of depressive symptoms or anxiety problems and the presence/absence of well-being or calmness. These findings indicate that interventions which measured change using the CES-D or STAI are already measuring the amelioration of depressive and anxiety symptoms by *simultaneously* reducing negative constructs and increasing positive constructs.

The second of PCP's predictions, that positive constructs provide important unshared explanatory power in relation to psychological problems (Johnson & Wood, 2017; Joseph & Wood, 2010; Wood & Johnson, 2016; Wood & Tarrier, 2010), does not seem particularly novel when considered beyond positive psychology in the broader context of psychological science. An easy illustration that this prediction is already well-established comes from the literature on stress, for example. Cohen and Wills' (1985) buffering hypothesis (now cited almost 14,000 times) posited that perceived social support buffers people against stress, reducing the negative consequences of stress. Probably hundreds of thousands of studies in the mental health field have examined the moderating and mediating role of a vast range of variables, including well-being variables.

PCP seems to be correct in asserting that some fields, such as positive psychology and mental health research, *tend* to have *focused* on particular variables and to have adopted particular zeitgeists. There certainly seems to be potential within both fields to more *explicitly* focus on exploring variables that have been understudied in that particular field, but these changes would appear to represent a shift in emphasis rather than something totally new. For this reason, it seems unlikely that PCP and "positive interventions" will revolutionise mental health treatments, somehow vastly improving on what decades of research, hundreds of therapy models, and billions of dollars of funding have attempted to achieve. However, PCP's call for integration may well serve to bring advances from disparate literatures together to the benefit of all.

Third, the PCP and positive psychology literatures need to be held to the same high standards as the existing treatment literature for psychological problems. For example, active treatments need to be pitted against one-another, comparing so-called "positive interventions" against psychological interventions that have an established evidence-base. Others have discussed this issue. For example, Wood et al. (2010) raised concerns about the control groups used in trials of gratitude interventions (a "positive" intervention), such as that in one study (Seligman et al., 2005), participants were recruited partially via a book detailing why

this intervention should work (potentially creating biased and overly positive self-reporting of outcome in the gratitude condition).

Finally, in their discussion of the potential value of "positive interventions" to the mental health field, Johnson and Wood (2017) state that these interventions "are usually 'skill building' treatments. As such, positive psychology interventions differ qualitatively from traditional interventions" (p. 341). However, one of the central purposes of many psychological therapies, such as cognitive behavioural therapy (Beck, 2011), dialectical behavioural therapy (Linehan, 1993), and schema therapy (numerous other therapies could be listed), is skilled building. For example, Beck (2011) stated "to effectively solve their problems, patients may need to learn new skills" (p. 296). Thus, overall, although PCP has made a valuable contribution by drawing attention to the need to integrate positive psychology with the existing mental health research literature, some of its more subsidiary statements may seem somewhat overstated.

6.6. Replication

Although this thesis was predominantly contextualised by PCP, it was also, to a lesser extent, contextualised by the methodological need for scientific findings to replicate. Chapters 2 and 3 provided direct replications of existing studies, which can be interpreted as evidence against the replication crisis. As discussed in the Introduction, "Exact" or "direct" replications are replications that operationalize both the independent and the dependent variable(s) in exactly the same way as the original study. The purpose of such replication attempts is to verify the original findings (Stroebe & Strack, 2014).

Chapters 4 and 5 described the development of the SABS and NSIBS; I demonstrated that important findings replicated here as well. The structural results observed in Samples 1 and 2 did not provide a definitive indication of the factor structure of cognitions about SA or cognitions about NSSI. This issue was addressed by collecting additional independent samples. EFA was conducted on each separate sample to locate the best fitting factor structure. The same factor structure emerged across Samples 3, 4, and 5, indicating that robust and replicable factor structures characterize SA and NSSI cognitions.

It is possible, although highly unlikely that the same EFA structural results emerged across Samples 3, 4, and 5 because of particular idiosyncrasies (e.g., random error) of these datasets and that findings may not generalise to other datasets (Brown, 2015). Without evidence of replication, it is possible that items may intercorrelate less strongly in a new sample or that the factor structure will vary as a function of sample differences (Smith & McCarthy, 1995). I therefore conducted a CFA on an independent sample (Sample 6). The

series of EFAs and a CFA across six independent, large samples meant that in the end I was able to hone in on a factor structure for the SABS and NSIBS that seem fairly robust and generalizable.

Replication also featured strongly during the validation of the SABS and NSIBS in that all of the psychometric analyses were conducted on at least two independent samples. Conclusions were then drawn based on the consensus of results across different samples, which again, provided some confidence that results are relatively robust and generalizable.

Revisiting the criteria required for a solid replication, power was adequate in each of the studies and the replications were conducted on samples from very different areas of the world. However, a downside of the replication presented in Chapter 2 is that it was conducted by the same research team (Wood and Taylor authored the 2010 paper on the CES-D that was replicated here).

6.6.1. The Replication Crisis

It is important for contemporary PhD theses to demonstrate an awareness of the replication crisis because this so-called "crisis" is a major topical debate currently facing all of psychological science. Moreover, Everett and Earp (2015) have argued that the PhD thesis is the ideal opportunity to contribute replication attempts to the field. The Introduction section (Chapter 1) provided an overview of the key tenants of replicability and the replication crisis in order to demonstrate my awareness of this literature.

Replication has historically been encouraged by some journals such as the *Journal of Personality and Social Psychology*, although it has often taken the form of excluding counter-explanations for results, rather than using different samples or research designs to test the same research question in different ways.

References

Achenbach, T. M. (1991). *Manual for the Youth Self-Report and 1991 profile*. Burlington, Vermont: University of Vermont, Department of Psychiatry.

Adams, J., Rodham, K., & Gavin, J. (2005). Investigating the "self" in deliberate self-harm. *Qualitative Health Research, 15,* 1293-1309.

Ajzen, I., & Driver, B. L. (1991). Prediction of leisure participation from behavioural, normative, and control beliefs: An application of the theory of planned behaviour. *Leisure Sciences, 13,* 185–204.

Ajzen, I., & Driver, B. L. (1992). Application of the theory of planned behaviour to leisure choice. *Journal of Leisure Research, 24,* 207–224.

American Psychological Association. Science paper shows low replicability of psychology studies; 2015. Accessed: 2015-10-31. http://bit.ly/1NogtsD

Amtmann, D., Kim, J., Chung, H., Bamer, A. M., Askew, R. L., Wu, S., Cook, K. F., & Johnson, K. L. (2014). Comparing CESD-10, PHQ-9, and PROMIS Depression Instruments in Individuals with Multiple Sclerosis. *Rehabilitation Psychology, 59,* 220–229.

Anderson, C. J., Bahnik, S., Barnett-Cowan, M., Bosco, F. A., Chandler, J.,… & Zuni, K. (2016). Response to comment on "Estimating the reprpducibility of psychological science. *Science, 351,* 1037.

Andrade, N. N., Hishinuma, E. S., McDermott, J. F., Jr., Johnson, R. C., Goebert, D. A., Makini, G. K., Jr., et al. (2006). The National Center on Indigenous Hawaiian Behavioral Health study of prevalence of psychiatric disorders in Native Hawaiian adolescents. *Journal of the American Academy of Child & Adolescent Psychiatry, 45,* 26–36.

Andresen, R., Oades, L., & Caputi, P. (2003). The experience of recovery from schizophrenia: Towards an empirically-validated stage model. *Australian and New Zealand Journal of Psychiatry, 37,* 586-594.

Anestis, M. D., Soberay, K. A., Gutierez, P. M., Hernandez, T. D., & Joiner, T. E. (2014). Reconsidering the link between impulsivity and suicidal behavior. *Personality and Social Psychology Review, 18,* 366-386.

Baker, M. (2016). Is there a reproducibility crisis? *Nature,* 533, 452-454.

Bakker, M., van Dijk, A., & Wicherts, J. M. (2012). The rules of the game called psychological science. *Perspectives on Psychological Science, 7,* 543–554.

Barnes, L. L. B., Harp, D., & Jung, W. S. (2002). Reliability generalization of scores on the Spielberger State–Trait Anxiety Inventory. *Educational and Psychological Measurement*, *62*, 603–618.

Baron-Cohen, S. (2012). Autism, empathizing-systematizing (E-S) theory, and pathological altruism (p 345-348). In: B, Oakley., A, Knafo., G, Madhava., & D. S, Wilson (Eds). *Pathological Altruism*. Oxford University Press, Oxford, UK.

Barrett, L. F., & Russell, J. A. (1998). Independence and bipolarity in the structure of current affect. *Journal of Personality and Social Psychology*, *74*, 967–984.

Batterham, P. J., Ftanou, M., Pirkis, J., Brewer, J. L., Mackinnon, A. J., Beautrais, A., Fairweather-Schmidt, A. K., & Christensen, H. (2014). A Systematic review and evaluation of measures for suicidal ideation and behaviors in population-based research. *Psychological Assessment*, *27*, 501-512.

Baumeister, R. F. (1990). Suicide as escape from self. *Psychological Review*, *97*, 90–113.

Baumeister, R. F., Stillwell, A. M., & Heatherton, T. F. (1994). Guilt: An interpersonal approach. *Psychological Bulletin*, *115*, 243-267.

Baumeister, R. F., Stillwell, A. M., & Heatherton, T. F. (1995). Personal narratives about guilt: Role in action control and interpersonal relationships. *Basic and Applied Social Psychology*, *17*, 173-198.

Baumeister, R. F., Vohs, K. D., DeWall, C. N., & Zhang, L. (2007). How emotion shapes behavior: Feedback, anticipation, and reflection, rather than direct causation. *Personality and Social Psychology Review*, *11*, 167-203.

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

Beck, J. S. (2007). *The Beck diet solution*. Birmingham, AL: Oxmoor House Publications.

Beck, A.T., Emery, G. & Greenberg, R.L. (1985). *Anxiety disorders and phobias: A cognitive perspective*. New York: Basic Books.

Beck, A.T., & Haigh, E.A.P. (2014). Advances in cognitive theory and therapy: The generic cognitive model. *Annual Review of Clinical Psychology*, *10*, 1-24.

Beck, A. T., Rush, A. J., Shaw, B. F., & Emery, G. (1979). *Cognitive therapy of depression*. New York: The Guilford Press.

Beck, A. T., & Steer, R. A. (1991). *Beck Scale for Suicidal Ideation: Manual*. New York, NY: Psychological Corporation.

Beck, A. T., & Steer, R. A. (1993). *Manual for Beck Hopelessness Scale*. San Antonio, TX: Psychological Corporation.

Beck, A. T., Weissman, A., Lester, D., & Trexler, L. (1974). The measurement of pessimism: The Hopelessness Scale. *Journal of Consulting and Clinical Psychology*, *42*, 861–865.

Beck, A. T., Wright, F. D., Newman, C. F., & Liese, B. S. (1993). *Cognitive therapy of substance abuse*. New York, NY: Guilford Press.

Bentall, R. (2003). *Madness explained: Psychosis and human nature*. The Penguin Press, London.

Bernstein, I. H., & Eveland, D. C. (1982). State vs. trait anxiety: A case study in confirmatory factor analysis. *Personality and Individual Differences, 3*, 361–372.

Bertamini, M., & Munafò, M. R. (2012). Bite-size science and its undesired side effects. *Perspectives on Psychological Science, 7*, 67–71.

Bodner, T. E. (2008). What improves with increased missing data imputations? *Structural Equation Modeling, 15*, 651-675.

Bolier, L., Haverman, M., Westerhof, G. J., Riper, H., Smit, F., & Bohlmeijer, E. (2013). Positive psychology interventions: A meta-analysis of randomized controlled studies. *BMC Public Health, 13*, 119.

Boeke, S., Duivenoorden, H., & Bonke, B. (1984). Agreement for surgical patients on two situations for the trait anxiety inventory. *Psychological Reports, 54*, 278.

Borenstein, M., Hedges, L. V., Higgins, J. P. T., & Rothstein, H. R. (2009). *Introduction to meta-analysis*. New York, NY: Wiley.

Bostwick, J. M., & Pankratz, V. S. (2000) Affective disorders and suicide risk: A reexamination. *American Journal of Psychiatry 2000, 157*, 1925-1932.

Bostic, T. J., Rubio, D. M., & Hood, M. (2000). A validation of the subjective vitality scale using structural equation modeling. *Social Indicators Research, 52*, 313-324.

Braver, S. L., Fogas, B., Sandler, I., & Volchik, S. (1986). *Development of a self-report aggressiveness scale for children: The Braver Aggressiveness Dimension Scale (BADS)*. Paper presented at the annual meeting of the Western Psychological Association, Seattle, Washington.

Braver, S. L., Thoemmes, F. J., & Rosenthal, R. (2014). Continuously cumulating meta-analysis and replicability. *Perspectives on Psychological Science, 9*, 333–342.

Briggs, S. R., & Cheek, J. M. (1986). The role of factor analysis in the development and evaluation of personality scales. *Journal of Personality, 54*, 106-148

Broadbent, D, E. (1965). A reformulation of the Yerkes-Dodson law. *British Journal of Mathematical and Statistical Psychology, 18*, 145–157.

Brown, G. (2001). *A review of suicide assessment scales for intervention research with adults and older adults*. Retrieved from http://www.suicidology.org/c/document_library/get_file?folderId=235&name=DLFE-113.pdf.

Brown, G. K., Henriques, G. R., Sosdjan, D., et al (2004) Suicide intent and accurate expectations of lethality: predictors of medical lethality of suicide attempts. *Journal of Consulting and Clinical Psychology, 72*, 1170 -1174.

Brown, T. B., & Kimball, T. (2013). Cutting to live: A phenomenology of self-harm. *Journal of Marital & Family Therapy, 39*, 195-208.

Bryan, C.J., Rudd, M.D., Wertenberger, E., Etienne, N., Ray-Sannerud, B.N., Morrow, C.E., Peterson, A.L., & Young-McCaughon, S. (2014). Improving the detection and prediction of suicidal behavior among military personnel by measuring suicidal beliefs: an evaluation of the Suicide Cognitions Scale. *Journal of Affective Disorders*, *159*, 15-22.

Butler, A. C., Chapman, J. E., Forman, E. M., & Beck, A. T. (2006). The empirical status of cognitive-behavioral therapy: A review of meta-analyses. *Clinical Psychology Review, 26*, 17-31.

Butler, G., Fennell, M., & Hackmann, A. (2010). *Cognitive behavioural therapy for anxiety disorders: Mastering clinical challenges*. New York: The Guilford Press.

Cacioppo, J. T., Gardner, W. L., & Berntson, G. G. (1999). The affect system has parallel and integrative processing components: Form follows function. *Journal of Personality and Social Psychology, 76*, 839–855.

Canals, J., Marti-Henneberg, C., Fernandez-Ballart, J., Clivillle, R., & Domenech, E. (1992). Scores on the State-Trait Anxiety Inventory for children in a longitudinal study of pubertal Spanish youth. *Psychological Reports, 71*, 503-512.

Carver, C. S. (2004). Negative affects deriving from the behavioral approach system. *Emotion, 4*, 3–22.

Carver, C. S., & Scheier, M. F. (1998). *On the self-regulation of behavior*. New York: Cambridge University Press.

Carver, C., Scheier, M. F., & Segersrom, S. C. (2010). Optimism. *Clinical Psychology Review, 30*, 879-889.

Cesario, J. (2014). Priming, replication, and the hardest science. *Perspectives on Psychological Science, 9*, 40-48.

Chapman, A. L., Gratz, K. L., Brown, M., (2006). Solving the puzzle of deliberate self-harm: The experiential avoidance model. *Behavior Research and Therapy*, *44*, 371-394.

Cheung, G. W., & Rensvold, R. B. (2002). Evaluating goodness-of-fit indexes for testing measurement invariance. *Structural Equation Modeling*, *9*, 233–255.

Chen, F., Curran, P. J., Bollen, K. A., Kirby, J., & Paxton, P. (2008). An empirical evaluation of the use of fixed cutoff points in RMSEA test statistic in structural equation models. *Sociological Methods & Research*, *36*, 462-494.

Cheung, G. W., & Rensvold, R. B. (2002). Evaluating goodness-of-fit indexes for testing measurement invariance. *Structural Equation Modeling*, *9*, 233–255.

Claes, L., & Vandereycken, W. (2007). The Self-Injury Questionnaire-Treatment Related (SIQ-TR): Construction, reliability, and validity in a sample of female eating disorder patients. In P. M. Goldfarb (Ed.), *Psychological tests and testing research trends*. New York: Nova Science Publishers.

Clark, L. A., & Watson, D. (1995). Constructing validity: Basic issues in objective scale development. *Psychological Assessment, 7*, 309 –319.

Clarke L., & Whittaker M. (1998) Self-mutilation: culture, contexts and nursing response. *Journal of Clinical Nursing 7*, 129–137.

Cohen, J. (1962). The statistical power of abnormal-social psychology research: A review. *Journal of Abnormal and Social Psychology, 65*, 145-153.

Cohen, J. (1998). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ, Erlbaum.

Cohen, S., Kamarck, T., & Mermelstein, R. (1983). A global measure of perceived stress. *Journal of Health and Social Behavior*, *24*, 385–396.

Cohen S., & Wills, T. A. (1985). Stress, social support, and the buffering hypothesis. *Psychological Bulletin*, *98*, 310–357.

Connelly, B. S., & Ones, D. S. (2010). An other perspective on personality: Meta-analytic integration of observers' accuracy and predictive validity. *Psychological Bulletin, 136*, 1092–1122.

Cooper, M.J., Wells, A. and Todd, G. (2004). A cognitive theory of bulimia nervosa. *British Journal of Clinical Psychology, 43*, 1-16.

Crawford, S. L., Tennstedt, S. L., & McKinlay, J. B. (1995). A comparison of anlaytic methods for non-random missingness of outcome data. *Journal of Clinical Epidemiology, 48*, 209-19.

Cumming, G. (2014). The new statistics: Why and how. *Psychological Science, 25*, 7-29.

Department of Health (2009). *New horizons: Towards a shared vision for mental health consultation*. London: Department of Health.

De Los Reyes, A., & Kazdin, A. E. (2005). Informant discrepancies in the assessment of childhood psychopathology: A critical review, theoretical framework, and recommendations for further study. *Psychological Bulletin, 131*, 483-509.

De Moore, G. M., & Robertson, A. R. (1999). Suicide attempts by firearms and by leaping from heights: A comparative study of survivors. *American Journal of Psychiatry, 9,* 1425-1431.

Diener, E., Emmons, R. A., Larsen, R. J., & Griffin, S. (1985). The satisfaction with life scale. *Journal of Personality Assessment, 49,* 71–75.

Donders, A. R. T., van der Heijden, G. J. M. G., Stijnen, T., & Moons, K. G. M. (2006). Review: A gentle introduction to imputation of missing values. *Journal of Clinical Epidemiology, 59,* 1087-1091.

Duckworth, A. L., Steen, T. A., & Seligman, M. E. P. (2005). Positive psychology in clinical practice. *Annual Review of Clinical Psychology, 1,* 629–651.

Earp, B. D. (2015). Does religion deserve a place in secular medicine? *Journal of Medical Ethics.* E-letter.doi:10.1136/medethics-2013-101776.Available online at: https://www.academia.edu/11118590/Does_religion_deserve_a_place_in_secular_medicine.

Earp, B. D., & Trafimow, D. (2015). Replication, falsification, and the crisis of confidence in social psychology. *Frontiers in Psychology, 6,* 621.

Ellis, T. E., & Rufino, K. A. (2015). A psychometric study of the Suicide Cognitions Scale with psychiatric inpatients. *Psychological Assessment, 27,* 82-89.

Enders, C. K. (2010). *Applied missing data analysis*. New York: The Guilford Press.

Endler, N., Cox, B., Parker, J., & Bagby, M. (1992). Self-reports of depression and state-trait anxiety: Evidence for differential assessment. *Journal of Personality and Social Psychology, 63,* 832–838.

Ensel, W. M. (1986). Measuring depression: The CES-D Scale. In N. Lin, A. Dean, & W. M, Ensel (Eds.), *Social support, life events, and depression* (pp. 51-70). New York: Academic Press

Etz, A., & Vandekerckhove, J. (2016). A bayesian perspective on the reproducibility project: Psychology. *PLoS ONE, 11,* 1-12.

Everett, J. A. C. & Earp, B. D. (2015). A tragedy of the (academic) commons: Interpreting the replication crisis in psychology as a social dilemma for early-career researchers. *Frontiers in Psychology, 6*, 1-4.

Fabrigar, L. R., Wegener, D. T., MacCallum, R. C., & Strahan, E. J. (1999). Evaluating the use of exploratory factor analysis in psychological research. *Psychological Methods, 4*, 272-299.

Favazza, A. R. (1989). Why patients mutilate themselves. *Hospital and Community Psychiatry, 40*, 137–145.

Feldman, L. A. (1993). Distinguishing depression and anxiety in self-report: Evidence from confirmatory factor analysis on nonclinical and clinical samples. *Journal of Consulting and Clinical Psychology, 61*, 631–638.

Fennell, M. J. V. (1997). Low self-esteem: a cognitive perspective. *Behavioural and Cognitive Psychotherapy, 25*, 1–25.

Ferguson, E. (2016). Empathy: The good, the bad, and the ugly. In: A. M. Wood & J. Johnson (Eds.), The Wiley handbook of positive clinical psychology. New York City: Wiley.

Field, A. P., & Lester, K. J. (2010). Is there room for 'development' in models of information processing biases to threat in children and adolescents? *Clinical Child and Family Psychology Review, 13*, 315-332.

Fitzsimons, G. M., Finkel, E. J., & vanDellen, M. R. (2015). Transactive goal dynamics. *Psychological Review, 122*, 648-673.

Fleiss, J. (1986). *The design and analysis of clinical experiments*. New York: Wiley.

Flett, G. L., Vredenburg, K., & Krames, L. (1997). The continuity of depression in clinical and nonclinical samples. *Psychological Bulletin, 121*, 395–416.

Flora, D. B., & Curran, P. J. (2004). An empirical evaluation of alternative methods of estimation for confirmatory factor analysis with ordinal data. *Psychological Methods, 9*, 466-491.

Fowler, J. C., Charak, R., Elhai, J. D., Allen, J. G., Frueh, B. C., & Oldham, J. M. (2014). Construct validity and factor structure of the difficulties in Emotion Regulation Scale among adults with severe mental illness. *Journal of Psychiatric Research*, 58, 175-180.

Gable, S., Haidt, J., 2005. What (and why) is positive psychology? *Review of General Psychology, 9*, 103–110.

Garrison, C. Z., Addy, C. L., Jackson, K. L., McKeown, R. E., & Walller, J. L. (1991). The CES-D as a screen for depression and other psychiatric disorders in adolescents. *Journal of the American Academy of Child & Adolescent Psychiatry, 30,* 636–641.

Gámez, W., Chmielewski, M., Kotov, R., Ruggero, C., & Watson, D. (2011). Development of a measure of experiential avoidance: The Multidimensional Experiential Avoidance Questionnaire. *Psychological Assessment, 23,* 692–713.

Gamez, W., Chmielewski, M., Kotov, R., Ruggero, C., Suzuki, N., & Watson, D. (2014). The Brief Experiential Avoidance Questionnaire: Development and initial validation. *Psychological Assessment, 26,* 35-45.

Gelman, A., & Loken, E. (2014). The statistical crisis in science. *American Scientist, 102,* 460–465.

Gilbert, P. (1998). Evolutionary psychopathology: Why isn't the mind better designed than it is? *British Journal of Medical Psychology, 71,* 353-373.

Gilbert, P., & Allan, S. (1998). The role of defeat and entrapment (arrested flight) in depression: An exploration of an evolutionary view. *Psychological Medicine, 28,* 585–598.

Gilbert, D. T., King, G., Pettigrew, S., & Wilson, T. D. (2016). Comment on "Estimating the reproducibility of psychological science. Science, 351, 1037.

Giner-Sorolla, R. (2012). Science or art? How aesthetic standards grease the way through the publication bottleneck but undermine science. *Perspectives on Psychological Science, 7,* 562–571.

Goldstein, R. B., Black, D. W., Nasrallah, A., & Winokur, G. (1991). The prediction of suicide: Sensitivty, specificity, and predictive value of a multivariate model applied to suicide among 1906 patients with affective disorders. *Archives of General Psychiatry, 48,* 418-422.

Gómez, O. S., Juzgado, N. J., & Vegas, S. (2010). Replications types in experimental disciplines. Proceedings of the International Symposium on Empirical Software Engineering and Measurement (Bolzano: ESEM).

Gottschall, A. C., West, S. G., & Enders, C. K. (2012). A Comparison of Item-Level and Scale-Level Multiple Imputation for Questionnaire Batteries. *Multivariate Behavioral Research, 47,* 1-25

Graham, J. W. (2009). Missing data analysis: Making it work in the real world. *Annual Review of Psychology, 60,* 549–576.

Graham, J. W., Olchowski, A. E., & Gilreath, T. D. (2007). How many imputations are really needed? Some practical clarifications of multiple imputation theory. *Prevention Science, 8*, 206--213.

Graham, J. W., & Schafer, J. L. (1999). On the performance of multiple imputation for multivariate data with small sample size. In: R. Hoyle, *Statistical strategies for small sample research* (pp. 1–29). Thousand Oaks, CA: Sage.

Gratz. K. L. (2001). Measurement of deliberate self-harm: preliminary data on the Deliberate Self-Harm Inventory. *Journal of Psychopathology and Behavioural Assessment*, *23*, 253–63.

Gratz, K. L., & Roemer, L. (2004). Multidimensional assessment of emotion regulation and dysregulation: Development, factor structure, and initial validation of the Difficulties in Emotion Regulation Scale. *Journal of Psychopathology and Behavioral Assessment*, *26*, 41-54. (Erratum published December 2008, *Journal of Psychopathology and Behavioral Assessment*, *30*, 315).

Gray, J. A. (1990). Brain systems that mediate both emotion and cognition. *Cognition and Emotion, 4*, 269–288.

Green, D. P., Goldman, S. L., & Salovey, P. (1993). Measurement error masks bipolarity in affect ratings. *Journal of Personality and Social Psychology*, *64*, 1029-1064.

Green, J.M., Statham, H. and Snowdon, C., *Cambridge Prenatal Screening Study: Main Study, 1990-1991* [computer file]. Colchester, Essex: UK Data Archive [distributor], February 1996. SN: 3462, http://dx.doi.org/10.5255/UKDA-SN-3462-1.

Gros, D. F., Antony, M. M., Simms, L. J., & McCabe, R. E. (2007). Psychometric properties of the state–trait inventory for cognitive and somatic anxiety (STICSA): Comparison to the State–Trait Anxiety Inventory (STAI). *Psychological Assessment*, 19, 369-381.

Guadagnoli, E. & Velicer, W. F. (1988). Relation of sample size to the stability of component patterns. *Psychological Bulletin*, *103*, 265-275.

Harris (2000). Self-harm: Cutting the bad out of me. *Qualitative Health Research, 10*, 164-173.

Hankin, B.L., Fraley, C.R., Lahey, B.B., & Waldman, I.D. (2005). Is depression best-viewed as a continuum or discrete category? A taxometric analysis of childhood and adolescent depression in a population-based sample. *Journal of Abnormal Psychology, 114*, 96-110.

Harvey, A., Watkins, E., Mansell W., & Shafran, R. (2004). *Cognitive behavioural processes across psychological disorders: A transdiagnostic approach to research and treatment*. Oxford University Press, Oxford.

Haslam, N., Holland, E., & Kuppens, P. (2012). Categories versus dimensions in personality and psychopathology: A quantitative review of taxometric research. *Psychological Medicine, 42*, 903-920.

Hawton, K., Rodham, K., Evans, E., Harriss, L. (2009) Adolescents who self harm: A comparison of those who go to hospital and those who do not. *Child and Adolescent Mental Health, 14*, 24-30.

Hayes, S. C., Strosahl, K. D., & Wilson, K. G. (1999). *Acceptance and commitment therapy: An experiential approach to behavior change*. New York: Guilford Press.

Held, B. S. (2016). Why clinical psychology should not go positive and/or negative. In: A. M. Wood & J. Johnson (Eds.), *The Wiley handbook of positive clinical psychology*. New York City: Wiley.

Held, B. S. (2017). Positive psychology's a prior problem. *Journal of Humanistic Psychology, 01,* 1-30.

Higgins, E. T. (1997). Beyond pleasure and pain. *American Psychologist, 52*, 1280–1300.

Hill, R. M., Rey, Y., Marin, C. E., Sharp, C., Green, K. L., & Pettit, J. W. (2014). Evaluating the Interpersonal Needs Questionnaire: Comparison of the reliability, factor structure, and predictive validity across five versions. *Suicide and Life-Threatening Behavior, 45*, 302-314.

Hishinuma, E. S., Miyamoto, R. H., Nishimura, S. T., Goebert, D. A., Yuen, N, Y. C., Makini, G. K., Andrade, N. N., Johnson, R. C., & Carlton, B. S. (2001). Prediction of anxiety disorders using the State-Trait Anxiety Inventory for multiethnic adolescents. Anxiety Disorders, 15, 511-533.

Hishinuma, E. S., Miyamoto, R. H., Nishimura, S. T., & Nahulu, L. B. (2000). Differences in State–Trait Anxiety Inventory scores for ethnically diverse adolescents in Hawaii. *Cultural Diversity and Ethnic Minority Psychology, 6*, 73– 83.

Hogan, R. T. (1983). A socioanalytic theory of personality. In M. Page (Ed.), *1982 Nebraska Symposium on Motivation* (pp. 55–89). Lincoln: University of Nebraska Press.

Holden, R. R., & DeLisle, M. M. (2005). Factor analysis of the Beck Scale for Suicide Ideation with female suicide attempters. *Assessment, 12* 231–8.

Holden, R. R., Kerr, P. S., Mendonca, J. D., & Velamoor, V. R. (1998). Are some motives more linked to suicide proneness than others? *Journal of Clinical Psychology, 54*, 569-576.

Hu, L-T., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structural analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling, 6*, 1-55.

Hunsley, J., & Meyer, G. J. (2003). The incremental validity of psychological testing and assessment: Conceptual, methodological, and statistical issues. *Psycholgoical Assessment, 15*, 446-55.

IBM Corp. Released 2012. *IBM SPSS Statistics for Windows, Version 21.0*. Armonk, NY: IBM Corp.

Ioannidis, J. P. (2005). Contradicted and initially stronger effects in highly cited clinical research. *Journal of the American Medical Association, 294*, 218–228.

Ivanoff, A., Jang, S., Smyth, N., & Linehan, M. (1994). Fewer reasons for staying alive when you are thinking of killing yourself: The Brief Reasons for Living Inventory. *Journal of Psychopathology and Behavioral Assessment, 16*, 1-13.

Ivtzan, I., Lomas, T., Hefferon, K., & Worth, P. (2015). *Second wave positive psychology: Embracing the dark side of life*. London, UK, Routledge.

John, L. K., Loewenstein, G., & Prelec, D. (2012). Measuring the prevalence of questionable research practices with incentives for truth telling. *Psychological Science, 23*, 524–532.

Joiner, T. (2005). *Why people die by suicide*. Cambridge, MA, US: Harvard University Press.

Joiner, T. E., Pfaff, J. J., & Acres, J. G. (2002). A brief screening tool for suicidal symptoms in adolescents and young adults in general health settings: reliability and validity data from the Australian National General Practice Youth Suicide Prevention Project. *Behaviour Research and Therapy, 40*, 471-81.

Johnson, J., & Wood, A. M. (2017). Integrating positive and clinical psychology: Viewing human functioning as continua from positive to negative can benefit clinical assessment, interventions and understandings of resilience. *Cognitive Therapy and Research, 41*, 335-349.

Joseph, S. (2006). Measurement in depression: Positive psychology and the statistical bipolarity of depression and happiness. *Measurement, 4*, 156–161.

Joseph, S. (2007). Is the CES-D a measure of happiness? *Psychotherapy and Psychosomatics, 76*, 60.

Joseph, S., & Wood, A. M. (2010). Assessment of positive functioning in clinical psychology: Theoretical and practical issues. *Clinical Psychology Review*, *30*, 830–838.

Kapur, N., Cooper, J., O'Connor, R.C., & Hawton, K. (2013). Attempted suicide versus non-suicidal self injury: New diagnosis, or false dichotomy? *British Journal of Psychiatry, 202*, 326-238.

Kelly, R., Mansell, W., & Wood, A. M., Alatiq, Y., Dodd, A., & Searson, R. (2011). Extreme positive and negative appraisals of activated states interact to discriminate bipolar disorder from unipolar depression and non-clinical controls. *Journal of Affective Disorders, 134*, 438-444.

Kendler, K. S., Prescott, C. A., Myers, J., & Neale, M. C. (2003). The structure of genetic and environmental risk factors for common psychiatric and substance use disorders in men and women. *Archives of General Psychiatry, 60*, 929–937.

Kessler, R. C., Berglund, P., Borges, G., Nock, M., Wang, P. S. (2005). Trends in suicide ideation, plans, gestures, and attempts in the United States 1990-92 to 2001-2003. *Journal of the American Medical Association, 293*, 2487-2495.

Keyes, C.L.M. (2003). Complete mental health: An agenda for the 21st century. In C.L.M. Keyes., J. Haidt (Eds.), *Flourishing: Positive Psychology and the life well-lived* (pp. 293–312). Washington, DC: American Psychological Association.

Keyes, C.L.M. (2007). Promoting and protecting mental health as flourishing: A complementary strategy for improving national mental health. *American Psychologist, 62*, 95-108.

Kirmayer, L. J., Young, A., & Hayton, B. C. (1995). The cultural context of anxiety disorders. *Psychiatric Clinics of North America, 18*, 503– 521.

Klein, O., Doyen, S., Leys, C., Magalhães de Saldanha da Gama, P. A., Miller, S., Questienne, L., Cleermans, A. (2012). Low hopes, high expectations: Expectancy effects and the replicability of behvioral experiments. *Perspectives on Psychological Science, 572-584.*

Kleindienst, N., Bohus, M., Ludascher, P., Limberger, M. F., Kuenkele, K., Ebner-Priemer, U. W., Chapman, A. L., Reicherzer, M., Stieglitz, R. D., & Schmahl, C. (2008). Motives for nonsuicidal self-injury among women with borderline personality disorder. *Journal of Nervous and Mental Disease, 196*, 230-236.

Klonsky, E. D., & Glenn, C. R. (2009). Assessing the functions of non-suicidal self-injury: Psychometric properties of the Inventory of Statements About Self-injury (ISAS). *Journal of Psychopathology and Behavioral Assessment*, *31*, 215-219.

Klonsky, E. D. & May, A. M. (2014). Differentiating suicide attempters from suicide ideators: A critical frontier for suicidology research. *Suicide and Life-Threatening Behavior*, *1*, 1-5.

Klonsky, E.D., May, A.M., & Glenn, C.R. (2013). The relationship between nonsuicidal self-injury and attempted suicide: Converging evidence from four samples. *Journal of Abnormal Psychology*, *122*, 231-237.

Knight, R. G., Waal-Manning, H. J., & Spears, G. F. (1983). Some norms and reliability data for the State-Trait Anxiety Inventory and the Zung Self-Rating Depression Inventory. *Cognitive Therapy and Research*, *22*, 245-249.

Kodaka, M., Postuvan, V., Inagaki, M., & Yamada, M. (2010). A systematic review of scales that measure attitudes toward suicide. *International Journal of Social Psychiatry*, *57*, 338-361.

Kroenke, K, Spitzer, R. L., Williams, J. B. W. (2001). The PHQ-9: Validity of a brief depression severity measure. *Journal of General Internal Medicine*, *16*, 606–13.

Kroenke, K., Spitzer, R. L., Williams, J. B. W., & Lowe, B. (2010). The Patient Health Questionnaire Somatic, Anxiety, and Depressive Symptom Scales: A systematic review. *General Hospital Psychiatry*, *32*, 345-359.

Kuppens, P., Tuerlinckx, F., Russell, J. A., & Barrett, L. F. (2013). The relation between valence and arousal in subjective experience. *Psychological Bulletin*, *139*, 917-940.

LeBel, E. P.,& Peters, K. R. (2011). Fearing the future of empirical psychology: Bem's (2011) evidence of psi as a case study of deficiencies in moral research practice. *Review of General Psychology*, *15*, 371-379.

Ledesma, R. D., & Valero-Mora, P. (2007). Determining the number of factors to retain in EFA: An easy-to-use computer program for carrying out parallel analysis. *Practical Assessment, Research & Evaluation*, *12*, 1-11.

Lewis, S. P. & Santor, D. A. (2008). Development and Validation of the Self-Harm Reasons Questionnaire. *Suicide & Life-threatening Behaviour*, *38*, 104-115.

Linehan, M. M. (1993). *Cognitive behavioral treatment of borderline personality disorder.* New York: The Guilford Press.

Linehan, M. M. & Comtois, K. (1996). *Lifetime parasuicide history*. University of Washington, Seattle, WA, Unpublished work.

Linehan, M. M., Comtois, K. A., Brown, M. Z., Heard, H. L., Wagner, A. (2006). Suicide Attempt Self-Injury Interview (SASII): Development, reliability, and validity of a scale to assess suicide attempts and intentional self-injury. *Psychological Assessment, 18*, 303–312.

Linehan, M. M., Goodstein, J. L., Nielsen, S. L., & Chiles, J. A. (1983). Reasons for staying alive when you are thinking of killing yourself: The reasons for living inventory. *Journal of Consulting and Clinical Psychology, 51*, 276-286.

Linley, P. A., Joseph, S., Harrington, S., & Wood, A. M. (2006). Positive psychology: Past, present, and (possible) future. *Journal of Positive Psychology, 1*, 3–16.

Lipsey, M.W., & Wilson, D.B. (2001). *Practical meta-analysis*. Thousand Oaks, CA: Sage.

Liu, G., & Gould, A. L. (2002). Comparison of alternative strategies for analysis of longitudinal trials with dropouts. *Journal of Biopharmaceutical Statistics, 12*, 207-26.

Lloyd, E., Kelley, M. L., & Hope, T. (1997, April). *Self-mutilation in a community sample of adolescents: Descriptive characteristics and provisional prevalence rates*. Paper presented at the Annual Meeting of the Society for Behavioral Medicine, New Orleans, Louisiana.

Loevinger, J. (1957). Objective tests as instruments of psychological theory. *Psychological Reports, 3*, 635-694.

Lomas, T. (2015). Positive cross-cultural psychology: Exploring similarity and difference in constructions and experiences of wellbeing. *International Journal of Wellbeing, 5*, 60-77.

Lonner, W. J., & Ibrahim, F. A. (1989). Assessment in cross-cultural counseling. In P. B. Pedersen, J. G. Draguns, W. J. Lonner, & J. E. Trimble (Eds.), *Counseling across cultures* (3rd ed.) (pp. 299– 333). Honolulu, HI: University of Hawaii Press.

Lyubomirsky, S. & Lepper, H. S. (1999). A measure of subjective happiness: Preliminary reliability and construct validation. *Social Indicators Research, 46*, 137-155.

Ma, A. M., & Klonsky, E. D. (2013). Assessing motivations for suicide attempts: Development and psychometric properties of the Inventory of Motivations for Suicide Attempts (IMSA). *Suicide and Life-Threatening Behavior, 43*, 532-546.

MacCallum, R. C., Zhang, S., Preacher, K. J., & Rucker, D. D. (2002). On the practice of dichotomization of quantitative variables. *Psychological Methods, 7, 19-40.*

Makel, M. C., Plucker, J. A., & Hegarty, B. (2012). Replications in psychology research: How often do they really occur? *Perspectives on Psychological Science, 7*, 537–542.

Makini, G. K., Andrade, N. N., Nahulu, L. B., Yuen, N., Yate, A., McDermott, J, F., Danko, G. P., Nordquist, C. R., Johnson, R., & Waldron, J. A. (1996). Psychiatric symptoms of Hawaiian adolescents. *Cultural Diversity and Mental Health, 2*, 183-191.

Manson, S. M. (1996). Culture and DSM-IV: Implications for the diagnosis of mood and anxiety disorders. In: J. E. Mezzich, A. Kleinman, H. Fabrega Jr., & D. L. Parron (Eds.), *Culture and psychiatric diagnosis: a DSM-IV perspective* (pp. 99– 113). Washington, DC: American Psychiatric Association.

Markon, K. E. (2010). How things fall apart: Understanding the nature of internalizing through its relationship with impairment. *Journal of Abnormal Psychology*, *119*, 447-458.

Markon, K. E., & Chmielewski, M., & Miller, C. J. (2011). The reliability and validity of discrete and continuous measures of psychopathology: A quantitative review. *Psychological Bulletin, 137*, 856-879.

Matson, J. L., Bamburg, J. W., Cherry, K. E., & Paclawskyj, T. R. (1999). A validity study on the Questions About Behavioral Function (QABF) Scale: Predicting treatment success for self-injury, aggression, and stereotypies. *Research in Developmental Disabilities, 20*, 163-75.

Maxwell, S. E., Lau, M. Y, & & Howard, G. S. (2015). Is psychology suffering from a replication crisis? What does "failure to replicate" really mean? *American Psychologist, 70*, 487-498.

Mazzuchelli, T. G., Kane, R., & Rees, C. S. (2010). Behavioral activation interventions for well-being: A meta-analysis. *The Journal of Positive Psychology*, *5*, 105-121.

McArdle, J.J., Johnson, R.C., Hishinuma, E.S., Miyamoto, R.H., & Andrade, N.N. (2001). Structural equation modeling of group differences in CES-D ratings of Native Hawaiian and non-Hawaiian high school students. *Journal of Adolescent Research*, *16*, 108-149.

McClelland, G. H., & Judd, C. M. (1993). Statistical difficulties of detecting interactions and moderator effects. *Psychological Bulletin, 114*, 376-390.

McNulty, J.K., & Fincham, F.D. (2012). Beyond positive psychology? Toward a contextual view of psychological processes and well-being. *American Psychologist*, *67*, 101-110.

Meade, A. W., Johnson, E. C., & Braddy, P. W. (2008). Power and sensitivity of alternative fit indices in tests of measurement invariance. *Journal of Applied Psychology, 93*, 568-592.

Meehl, P. E. (1990). Why summaries of research on psychological theories are often uninterpretable. *Psychological Inquiry*, *1*, 108–141.

Mental Health Commission of Canada (2009). *Toward recovery and well-being*. Calgary: Mental Health Commission of Canada.

Metalsky, G. I., & Joiner, T. E. (1997). The hopelessness depression symptom questionnaire. *Cognitive Therapy and Research, 21, 359-384.*

Miller, G. (1990). *The SASSI adolescent manual*. Spencer, IN: The Spencer Evening World.

Miller, W. R., & Rollnick, S. (2002). *Motivational interviewing: Preparing people for change* (2nd ed.). New York: Guilford.

Mineka, S., Watson, D., & Clark, L. A. (1998). Comorbidity of anxiety and unipolar mood disorders. *Annual Review of Psychology, 49*, 377–412.

Mook, J., Van der Ploeg, H. M., & Kleijn, W. C. (1992). Symptom-positive and symptom-negative items in the State-Trait Anxiety Inventory: A comparison and replication. *Anxiety, Stress, and Coping*, *5*, 113–123.

Morgan, G. B., Hodge, K. J., Wells, K. E., & Watkins, M. W. (2015). Are fit indices biased in favor of bi-factor models in cognitive ability research?: A comparison of fit in correlated factors, higher-order, and bi-factor models via monte carlo simulations. *Journal of Intelligence*, *3*, 2-20.

Muehlenkamp, J. (2014). Distinguishing between suicidal and nonsuicidal self-injury. In M. Nock, *The Oxford handbook of suicide and self-injury* (pp. 23-46). Oxford: Oxford University Press.

National Institute for Health and Care Excellence (2011). *Self-harm: Longer-term management.* London: NICE.

Neuliep, J. W., & Crandall, R. (1990). Editorial bias against replication research. In J. W. Neuliep (Ed.), *Replication research in the social sciences* (pp. 85–90). London, England: SAGE.

Neuliep, J. W., & Crandall, R. (1993). Reviewer bias against replication research. *Journal of Social Behavior & Personality, 8*, 1–8.

Newman, D.L., Moffitt, T.E., Caspi, A. & Silva, P.A. (1998). Comorbid mental disorders: Implications for treatment and sample selection. *Journal of Abnormal Psychology, 107*, 305–311.

Nishimura, S. T., Hishinuma, E. S., Miyamoto, R. H., Goebert, D. A., Johnson, R. C., Yuen, N. Y. C., & Andrade, N. N., (2001). Prediction of DISC substance abuse and

dependency for ethnically diverse adolescents. *Journal of Substance Abuse, 13*, 597–607.

Nock, M. K., Borges, G., Bromet, E. J., Alonso, J., Angermeyer, M., Beautrais, A., et al. (2008a). Cross-national prevalence and risk factors for suicidal ideation, plans, and attempts. *The British Journal of Psychiatry, 192*, 98–105.

Nock, M. K., Holmberg, E. B., Photos, V. I., & Michel, B. D. (2007). The Self-Injurious Thoughts and Behaviors Interview: Development, reliability, and validity in an adolescent sample. *Psychological Assessment, 19*, 309–317.

Nock, M. K., Joiner, T. E., Gordon, K., Lloyd-Richardson, E., & Prinstein, M. J. (2006). Non-suicidal self-injury among adolescents: Diagnostic correlates and relation to suicide attempts. *Psychiatry Research, 144*, 65-72.

Nock, M. K., Prinstein, M. J., & Sterba, S. (2009). Revealing the form and function of self-injurious thoughts and behaviors: A real-time ecological assessment study among adolescents and young adults. *Journal of Abnormal Psychology, 118*, 816-827.

Nock, M. K., Wedig, M. M., Holmberg, E. B., & Hooley, J. M. (2008). The Emotion Reactivity Scale: Development, evaluation, and relation to self-injurious thoughts and behaviors. *Behavior Therapy, 39*, 107-116.

Novick, M. R., & Lewis, C. L. (1967). Coefficient alpha and the reliability of composite measurements. *Psychometrika, 32*, 1-13.

Nunnally, J. (1978). *Psychometric theory* (2nd ed.). New York: McGraw-Hill.

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

O'Connor, R. C. (2011). The integrated motivational-volitional model of suicidal behaviour. *Crisis, 32*, 295-298.

O'Connor, R. C., Nock, M. K. (2014). The psychology of suicidal behaviour. *Lancet Psychiatry, 1*, 73-85.

Open Science Collaboration. (2012). An open, large-scale, collaborative effort to estimate the reproducibility of psychological science. *Perspectives on Psychological Science, 7*, 657–660.

Open Science Collaboration. (2015). Estimating the reproducibility of psychological science. *Science, 349*, 6521.

Osman, A., Bagge, C. L., Gutierrez, P. M., Konick, L. C. Kopper, B. A., & Barrios, F. X. (2001). The Suicidal Behaviors Questionnaire-Revised (SBQ-R): Validation with clinical and nonclinical samples. *Assessment, 8*, 443-454.

Osman, A., Gutierrez, P. M., Kopper, B. A., Barrios, F. X., & Chiros, C. E. (1998). The Positive and Negative Suicide Ideation Scale: Development and validation. *Psychological Reports*, *28*, 783-793.

Osman, A., Lamis, D. A., Freedenthal, S., Guttierrez, P. M., & McNaughton-Cassill, M. (2014). The Multidimensional scale of perceived social support: Analyses of internal reliability, measurement invariance, and correlates across gender. *Journal of Personality Assessment*, *96*, 103-112.

Osuch, E. A., Noll, J. G., &Putnam, F. W. (1999). The motivations for self-injury in psychiatric inpatients. *Psychiatry*, *62*, 334–346.

Parrott, W. G. (2014). *The positive side of negative emotions*. New York, NY: Guilford.

Patil, P., Peng, R. D., & Leek, J. T. (2016). What should researchers expect when they replicate studies? A statistical view of replicability in psychological science. *Perspectives on Psychological Science*, *11*, 539-544.

Pearson, D. G., Deeprose, C., Wallace-Hadrill, S. M. A., Burnett Heyes, S., & Holmes, E. A. (2013). Assessing mental imagery in clinical psychology: A review of imagery measures and a guiding framework. *Clinical Psychology Review*, *33*, 1-23.

Peterson, C. (2006a). Course syllabus for special seminar in psychology: Positive psychology. Retrieved online from http://www.ppc.sas.upenn.edu/ppsyllabus2006peterson.pdf.

Peterson, C. (2006b). *A primer in positive psychology*. New York, Oxford University Press.

Piazza, N. J. (1996). Dual diagnosis and adolescent psychiatric inpatients. *Substance Use and Misuse*, *31*, 215–223.

Posner, K., Brodsky, B., Yershova, K., Buchanan, J., & Mann, J. (2014). In M. Nock, *The Oxford handbook of suicide and self-injury* (pp. 23-46). Oxford: Oxford University Press.

Powers, W. T. (2005). *Behaviour: The control of perception (*2nd ed.). New Canaan, CT, USA: Benchmark Publications.

Prescott, C.A., McArdle, J.J., Hishinuma, E.S., Johnson, R.C., Miyamoto, R.H., Andrade, N.N., Edman, J.L., Makini, G.K., Nahulu, L.B., Yuen, N.Y.C., & Carlton, B.S. (1998). Prediction of major depression and dysthymia from CES-D scores among ethnic minority adolescents. *Journal of the American Academy of Child and Adolescent Psychiatry*, *37*, 495-503.

R Development Core Team (2009). *R Project for Statistical Computing*. Vienna, Austria.

Radloff, L. S. (1977). The CES-D scale: A self-report depression scale for research in the general population. *Applied Psychological Measurement, 1*, 385–401.

Risberg, R. A., Stevens, M. J., & Graybill, D. F. (1995). Validating the adolescent form of the Substance Abuse Subtle Screening Inventory. *Journal of Child and Adolescent Substance Abuse, 4*, 25–41.

Roberts, R. E. (1980). Reliability of the CES-D Scale in different ethnic contexts. *Psychiatry Research, 2*, 125-134.

Rosenthal, R. (1979). The file drawer problem and tolerance for null results. *Psychological Bulletin, 86,* 638–641.

Rosnow, R. L., & Rosenthal, R. (1989). Statistical procedures and the justification of knowledge in psychological science. *American Psychologist, 44*, 1276-1284.

Rosseel, Y. (2012). Lavaan: an R package for structural equation modeling. *Journal of Statistical Software, 48*, 1–36.

Rothberg, J. M., & Geer-Williams, C. (1992). A comparison and review of suicide prediction scales. In R. W. Maris, A. L. Berman, J. T. Maltsberger, & R. I. Yufits (Eds.), *Assessment and prediction of suicide* (pp. 202-217). New York: The Guilford Press.

Rudd, M. D. (2000). The suicidal mode: A cognitive behavioral model of suicidality. *Suicide and Life-Threatening Behavior, 30,* 18–33.

Rudd, M. D. (2006). Fluid vulnerability theory: A cognitive approach to understanding the process of acute and chronic suicide risk. In T. E. Ellis (Ed.), *Cognition and suicide: Theory, research, and therapy* (pp. 355–368). Washington, DC: American Psychological Association.

Rudd, M. D., Joiner, T., & Rajab, M. H. (1996). Relationships among suicide ideators, attempters, and multiple attempters in a young-adult sample. *Journal of Abnormal Psychology, 105*, 541–550.

Rudd, M. D., Schmitz, W., McClenen, R., Joiner, T., Elkins, G., & Claassen, C. (in preparation). The Suicide Cognitions Scale: A suicide-specific measure of hopelessness.

Russell, J. A. (1980). A circumplex model of affect. *Journal of Personality and Social Psychology, 39*, 1161–1178.

Russell, J. A., & Carroll, J. M. (1999a). On the bipolarity of positive and negative affect. *Psychological Bulletin, 125*, 3-30.

Russell, J. A., & Carroll, J. M. (1999b). The phoenix of bipolarity: Reply to Watson and Tellegen (1999). *Psychological Bulletin, 125*, 611-617.

Ryff, C. D. (1989). Happiness is everything, or is it? Explorations on the meaning of psychological well-being. *Journal of Personality and Social Psychology*, *57*, 1069–1081.

Ryff, C. D., & Singer, B. (2003). Ironies of the human condition: Wellbeing and health on the way to mortality. In L. G. Aspinwall, & U. M. Staudinger (Eds.), *A psychology of human strengths* (pp. 271–287). Washington, DC: American Psychological Association.

Salehi, B., Cordero, M. I., & Sandi, C. (2010). Learning under stress: The inverted-U-shape function revisited. *Learning & Memory*, *17*, 522-30.

Sansone, R. A., Wiederman, M. W., & Sansone, L. A. (1998). The Self-Harm Inventory (SHI): Development of a scale for identifying self-destructive behaviors and borderline personality disorder. *Journal of Clinical Psychology, 54*, 973-983.

Santa Mina, E. E., Gallop, R., Links, P., Heslegrave, R., Pringle, D., Wekerle, C., & Grewal, P. (2006) The Self-Injury Questionnaire: The evaluation of the psychometric properties in a clinical population. *The Journal of Psychiatric and Mental Health Nursing*, *13*, 221-227.

Santor, D. A., Gregus, M., & Welch, A. (2006). Eight decades of measurement in depression. *Measurement*, *4*, 135-155.

Schafer, J. L., & Graham, J. W. (2002). Missing data: Our view of the state of the art. *Psychological Methods*, *7*, 147–77.

Scheier, M. F., Carver, C. S., & Bridges, M. W. (1994). Distinguishing optimism from neuroticism (and trait anxiety, self-mastery, and self-esteem): A reevaluation of the Life Orientation Test. *Journal of Personality and Social Psychology*, *67*, 1063−1078.

Schmidt, S. (2009). Shall we really do it again? The powerful concept of replication is neglected in the social sciences. *Review of General Psychology, 13*, 90–100.

Schmitt, N., & Stults, D.M. (1985). Factors defined by negatively keyed items: The result of careless respondents? *Applied Psychological Measurement*, *9*, 367–373.

Sedlmeier, P., & Gigerenzer, G. (1989). Do studies of statistical power have an effect on the power of studies? *Psychological Bulletin, 105*, 309-316.

Selby, E. A., Joiner, T. E., & Ribeiro, J. D. (2014). Comprehensive theories of suicidal behaviors. In M. Nock, *The Oxford handbook of suicide and self-injury* (pp. 286-307). Oxford: Oxford University Press. Siddaway, A. P., Taylor, P. J., Wood, A. M., & Schulz, J. (2015). A meta-analysis of the role of perceptions of defeat and entrapment

in depression, anxiety problems, posttraumatic stress disorder and suicidality. *Journal of Affective Disorders, 184,* 149-159.

Seligman, M.E.P., & Csikszentmihalyi, M. (2000). Positive psychology: An introduction. *American Psychologist, 55*, 5–14.

Shean, G., & Baldwin, G. (2008). Sensitivity and specificity of depression questionnaires in a college-age sample. *Journal of Genetic Psychology, 169,* 281-288.

Sheldon, K., Kashdan, T., & Steger, M. (Eds.). (2011). *Designing the future of positive psychology: Taking stock and moving forward*. New York: Oxford University Press.

Shrive, F. M., Stuart, H., & Quan, H. (2006). Dealing with missing data in a multi-question depression scale: A comparison of imputation methods. *BMC Medical Research Methodology, 6,* 57-67.

Siddaway, A.P., Taylor P.J., & Wood, A.M. (in press). Re-conceptualizing anxiety as a continuum that ranges from high calmness to high anxiety: The joint importance of reducing distress and increasing well-being. *Journal of Personality and Social Psychology*.

Siddaway, A. P., Wood, A. M., O'Carroll, R. E., & O'Connor, R. C. (in submission). Characterizing self-injurious cognitions Part I: Development of the Suicide Attempt Beliefs Scale (SABS) and the Nonsuicidal Self-Injury Beliefs Scale (NSIBS).

Siddaway, A. P., Wood, A. M., O'Carroll, R. E., & O'Connor, R. C. (in submission). Characterizing self-injurious cognitions Part II: Validation of the Suicide Attempt Beliefs Scale (SABS) and the Nonsuicidal Self-Injury Beliefs Scale (NSIBS).

Silverman, M. M. (2016). Challenges to defining and classifying suicide and suicidal behaviors (pp.11-35). In R. C. O'Connor & J. Pirkis, (Eds.) (2016). *The International Handbook of Suicide Prevention* (2nd ed.). Chichester: Wiley-Blackwell.

Silverman, M. M., Berman, A. L., Sanddal, N. D., O'Carroll, P. W., & Joiner, T. E. (2007a). Rebuilding the Tower of Babel: A revised nomenclature for the study of suicide and suicidal behaviors part 1: Background, rationale, and methodology. *Suicide and Life-Threatening Behavior, 37,* 248-263.

Silverman, M. M., Berman, A. L., Sanddal, N. D., O'Carroll, P. W., & Joiner, T. E. (2007b). Rebuilding the Tower of Babel: A revised nomenclature for the study of suicide and suicidal behaviors part 2: Suicide-related ideations, communications, and behaviors. *Suicide and Life-Threatening Behavior, 37,* 264-277.

Simmons, J. P., Nelson, L. D., & Simonsohn, U. (2011). Falsepositive psychology: Undisclosed flexibility in data collection and analysis allows presenting anything as significant. *Psychological Science, 22*, 1359–1366.

Sloman, L. (2000). How the involuntary defeat strategy relates to depression. In L. Sloman & P. Gilbert (Eds.), *Subordination and defeat: An evolutionary approach to mood disorders and their therapy* (pp. 47–67). Mahwah, NJ: Erlbaum.

Sloman, L., Gilbert, P., & Hasey, G. (2003). Evolved mechanisms in depression: The role and interaction of attachment and social rank in depression. *Journal of Affective Disorders, 74*, 107–121.

Snook, S. C., & Gorsuch, R. L. (1989). Component analysis versus common factor analysis: A Monte Carlo study. *Psychological Bulletin*, *106*, 148-154.

Snyder, C. R., & Lopez, S. J. (2007). *Positive psychology: The scientific and practical explorations of human strengths*. Thousand Oaks, CA: Sage.

Spielberger, C. D. (1989). *State–Trait Anxiety Inventory: a comprehensive bibliography* (2nd ed.). Palo Alto, CA: Consulting Psychologists Press.

Spielberger, C. D., & Diaz-Guerrero, R. (Eds.) (1986). *Cross-cultural anxiety, vol. 3*. Washington, DC: Hemisphere/Wiley.

Spielberger, C. D., Gorsuch, R. L., & Lushene, R. E. (1970). *State–Trait Anxiety Inventory for adults (Form X)*. Palo Alto, CA: Consulting Psychologists Press.

Spielberger, C. D., Vagg, P. R., Barker, L. R., Donham, G. W., & Westberry, L. G. (1980). The factor structure of the State–Trait Anxiety Inventory. In I. G. Sarason & C. D. Spielberger (Eds.), *Stress and anxiety, vol. 7* (pp. 95– 109). New York: Hemisphere/Wiley.

Spitzer, R. L., Kroenke, K., Williams, J. B. W., & Lowe, B. (2006). A brief measure for assessing generalized anxiety disorder — the GAD-7. *Archices of Internal Medicine, 166*, 1092–7.

Spitzer, R. L., Kroenke, K., & Williams, J. B. W. (1999). Validation and utility of a self-report version of PRIME-MD: The PHQ primary care study. *JAMA, 282*, 1737–44.

Steiger, J. H. (1980). Tests for comparing elements of a correlation matrix. *Psychological Bulletin, 87*, 245-251.

Stroebe, W., & Strack, F. (2014). The alleged crisis and the illusion of exact replication. *Perspectives on Psychological Science, 9*, 59-71.

Suzuki, T., Tsukamoto, K., & Abe, K. (2000). Characteristic factor structures of the Japanese version of the State-Trait Anxiety Inventory: Coexistence of positive-negative and State-Trait factor structures. *Journal of Personality Assessment, 74*, 447–458.

Tabachnick, B. G., & Fidell, L. S. (2007). *Using multivariate statistics* (6th ed.). Boston: Pearson/Allyn & Bacon. Chicago.

Tang, L., Song, J., Belin, T. R., & Unutzer, J. (2005). A comparison of imputation methods in a longitudinal randomized clinical trial. *Statistics in Medicine*, *24*, 2111-28.

Ten Have, M., de Graaf, R., van Dorsselaer, S., Verdurmen, J., van't Land, H., Vellebergh, W., et al. (2009). Incidence and course of suicidal ideation and suicide attempts in the general population. *Canadian Journal of Psychiatry, 54*, 824–833.

The British Psychological Society. (2010). *Division of Clinical Psychology: The core purpose and philosophy of the profession*. Leicester, UK: The British Psychological Society.

Turner, B. J., Chapman, A. L., & Gratz, K. L. (2014). Why stop self-injuring? Development of the reasons to stop self-injury questionnaire. *Behavior Modification, 38,* 69-106.

Valentine, J. C. Biglan, A., Boruch, R. F., Castro, F. G., Collins, L. M., Flay, B. R., et al. (2011). Replication in prevention science. *Prevention Science, 12*, 103-117.

Van Orden, K. A., Cukrowicz, K. C., Witte, T. K., & Joiner, T. E. (2012). Thwarted belongingness and perceived burdensomeness: Construct validity and psychometric properties of the Interpersonal Needs Questionnaire. *Psychological Assessment, 24,* 197-215.

Van Orden, K. A., Witte, T. K., Cukrowicz, K. C., Braithwaite, S. R., Selby, E. A., & Joiner, T. E. (2010). The interpersonal theory of suicide. *Psychological Review, 117*, 575-600.

Vautier, S., & Pohl, S. (2009). Do balanced scales assess bipolar constructs? The case of the STAI scales. *Psychological Assessment, 21*, 187-193.

Velicer, W. F. (1976). Determining the number of components from the matrix of partial correlations. *Psychometrika*, *41*, 321-327.

Velicer, W. F., Eaton, C. A., & Fava, J. L. (2000). Construct explication through factor or component analysis: A review and evaluation of alternative procedures for determining the number of factors or components. In R. D. Goffin & E. Helmes (Eds.), *Problems and solutions in human assessment: Honoring Douglas Jackson at seventy* (pp. 41-71). Boston, MA: Kluwer.

Vigneau, F., & Cormier, S. (2008). The factor structure of the State-Trait Anxiety Inventory: An alternative view. *Journal of Personality Assessment, 90*, 280–285.

Vagg, P. R., Spielberger, C. D., & O'Hearn, T. P. J. R. (1980). Is the State–Trait anxiety inventory multidimensional? *Personality and Individual Differences, 1*, 207-14.

Waller, N.G., & Meehl, P.E. (1998). *Multivariate taxometric procedures: Distinguishing types of continua*. Sage, Thousand Oaks, CA.

Warttig, S. L., Forshaw, M. J., South, J., White, A. K. (2013). New, normative, English-sample data for the Short Form Perceived Stress Scale (PSS-4). *Journal of Health Psychology, 18,* 1617-28.

Watson, D. (2009). Differentiating the mood and anxiety disorders: A quadripartite model. *Annual Review of Clinical Psychology, 5*, 221-247.

Watson, D., O'Hara, M. W., Chmielewski, M., McDade-Montez, E. A., Koffel, E., Naragon, K., & Stuart, S. (2008). Further validation of the IDAS: Evidence of convergent, discriminant, criterion, and incremental validity. *Psychological Assessment, 20*, 248-259.

Watson, D., O'Hara, M. W., Simms, L. J., Kotov, R., Chmielewski, M., McDade-Montez, E., . . . Stuart, S. (2007). Development and validation of the Inventory of Depression and Anxiety Symptoms (IDAS). *Psychological Assessment, 19,* 253-268.

Watson, D., & Tellegen, A. (1999). Issues in the dimensional structure of affect—Effects of descriptors, measurement error, and response formats: Comment on Russell and Carroll (1999). *Psychological Bulletin, 125,* 601-610.

Watson, D., Wiese, D., Vaidya, J., & Tellegen, A. (1999). The two general activation systems of affect: Structural findings, evolutionary considerations, and psychobiological evidence. *Journal of Personality and Social Psychology, 76*, 820–838.

Weiss, N. H., Gratz, K. L., & Lavender, J.M. (2015). Factor structure and initial validation of a multidimensional measure of difficulties in the regulation of positive emotions: The DERS-Positive. *Behavior Modification, 39*, 431-453.

Wells, A. (1997). *Cognitive therapy of anxiety disorders*. Chichester: Wiley.

Wells, A. (2000). *Emotional disorders and metacognition: Innovative cognitive therapy*. Chichester: Wiley.

White, I. R., Royston, P., & Wood, A. M. (2011). Multiple imputation using chained equations: Issues and guidance for practice. *Statistics in Medicine, 30,* 377-99.

Whitlock, J. L., Exner-Cortens, D., & Purington, A. (2014). Assessment of nonsuicidal self-injury: Development and initial validation of the Non-Suicidal Self-Injury-Assessment Tool (NSSI-AT). *Psychological Assessment, 26,* 935-946.

Williams, J. M. G. (1997). *Cry of pain*. London: Penguin.

Winter, D. A. (2003). Repertory grid technique as a psychotherapy research measure. *Psychotherapy research, 13,* 25-42.

Winters, N. C., Myers, K., & Proud, L. (2002). Ten-year review of rating scales. III: Scales assessing suicidality, cognitive style, and self-esteem. *Journal of the American Academy of Child and Adolescent Psychiatry, 41,* 1150-1181.

Witte, T. K., Joiner, T. E., Brown, G. K., Beck, A. T., Beckman, A., Duberstein, P., & Conwell, Y. (2006). Factors of suicide ideation and their relation to clinical and other indicators in older adults. *Journal of Affective Disorders, 94,* 165-172.

Wong, P.T.P. (2011). Positive psychology 2.0: Towards a balanced interactive model of the good life. *Canadian Psychology, 52,* 69-81.

Wood, A. M. & Johnson, J. (2016). *The Wiley handbook of Positive Clinical Psychology: An integrative approach to studying and improving well-being*. Chichester, UK: John Wiley & Sons Ltd.

Wood, A. M., & Tarrier, N. (2010). Positive Clinical Psychology: A new vision and strategy for integrated research and practice. *Clinical Psychology Review, 30,* 819-829.

Wood, A.M., Taylor, P.J. & Joseph, S. (2010). Does the CES-D measure a continuum from depression to happiness? Comparing substantive and artifactual models. *Psychiatry Research, 177,* 120–123.

Wood, J. V., Perunovic, W. Q., & Lee, J. W. (2009). Positive self-statements: Power for some, peril for others. *Psychological Science, 20,* 860-866.

Woods, C. M. (2006). Careless responding to reverse-worded items: Implications for confirmatory factor analysis. *Journal of Psychopathology and Behavioral Assessment, 28,* 189-194.

World Health Organization (2004). *Promoting mental health: Concepts, emerging evidence, practice*. Geneva: World Health Organization.

Yerkes, R. M., & Dodson, J. D. (1908). The relation of strength of stimulus to rapidity of habit-formation. *Journal of Comparative Neurology, 18,* 459–482.

Yik, M., Russell, J. A., & Steiger, J. H. (2011). A 12-point circumplex model of core affect. *Emotion, 11,* 705-731.

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

Zanarini, M. C., Vujanovic, A., Parachini, E. A., Boulanger, J. L., Frankenburg, F. R., & Hennen, J. (2003). A screening measure for BPD: The McLean Screening Instrument for Borderline Personality Disorder. *Journal of Personality Disorders, 17*, 568-573.

Zimet, G. D., Dahlem, N. W., Zimet, S. G., & Farley, G. K. (1988). The multidimensional scale of perceived social support. *Journal of Personality Assessment, 52*, 30-41.

Zimet, G. D., Powell, S. S., Farley, G. K., Werkman, S., & Berkoff, K. A. (1990). Psychometric characteristics of the multidimensional scale of perceived social support. *Journal of Personality Assessment, 55*, 610-617.

Zwick, W. R., & Velicer, W. F. (1986). Comparison of five rules for determining the number of components to retain. *Psychological Bulletin, 99*, 432–442.

APPENDIX

Factor Loadings for Chapter 2 Confirmatory Factor Analyses

Table A1. Factor Loadings for ML CFA using Hawai‘i sample, Single Factor Model

Latent Variables:

	Estimate	Std.Err	z-value	P(>|z|)	Std.lv	Std.all
F1 =~						
dep_1	0.502	0.014	36.009	0.000	0.502	0.573
dep_2	0.363	0.014	26.846	0.000	0.363	0.445
dep_3	0.699	0.014	49.943	0.000	0.699	0.739
dep_4	0.316	0.019	17.061	0.000	0.316	0.292
dep_5	0.556	0.016	33.797	0.000	0.556	0.544
dep_6	0.822	0.014	57.082	0.000	0.822	0.810
dep_7	0.042	0.018	2.340	0.019	0.042	0.041
dep_8	0.259	0.018	14.350	0.000	0.259	0.247
dep_9	0.532	0.012	43.579	0.000	0.532	0.668
dep_10	0.456	0.012	36.741	0.000	0.456	0.583
dep_11	0.497	0.016	31.161	0.000	0.497	0.507
dep_12	0.534	0.015	35.598	0.000	0.534	0.568
dep_13	0.393	0.015	26.794	0.000	0.393	0.444
dep_14	0.706	0.014	49.424	0.000	0.706	0.733
dep_15	0.325	0.012	26.418	0.000	0.325	0.439
dep_16	0.554	0.016	34.046	0.000	0.554	0.547
dep_17	0.509	0.012	41.769	0.000	0.509	0.646
dep_18	0.758	0.013	57.436	0.000	0.758	0.813
dep_19	0.518	0.013	39.192	0.000	0.518	0.614
dep_20	0.588	0.014	41.512	0.000	0.588	0.643

Table A2. Factor Loadings for ML CFA using Hawai‘i sample, Two Factor Model

Latent Variables:

	Estimate	Std.Err	z-value	P(>|z|)	Std.lv	Std.all
depression =~						
dep_1	0.505	0.014	36.197	0.000	0.505	0.576
dep_2	0.364	0.014	26.910	0.000	0.364	0.446
dep_3	0.699	0.014	49.959	0.000	0.699	0.739
dep_5	0.559	0.016	34.029	0.000	0.559	0.548
dep_6	0.822	0.014	57.107	0.000	0.822	0.810
dep_7	0.060	0.018	3.374	0.001	0.060	0.059
dep_9	0.528	0.012	43.129	0.000	0.528	0.663
dep_10	0.458	0.012	36.975	0.000	0.458	0.587
dep_11	0.497	0.016	31.167	0.000	0.497	0.508
dep_13	0.398	0.015	27.139	0.000	0.398	0.450
dep_14	0.711	0.014	49.817	0.000	0.711	0.738
dep_15	0.328	0.012	26.610	0.000	0.328	0.442
dep_17	0.514	0.012	42.188	0.000	0.514	0.652
dep_18	0.762	0.013	57.900	0.000	0.762	0.818
dep_19	0.520	0.013	39.373	0.000	0.520	0.617
dep_20	0.590	0.014	41.634	0.000	0.590	0.645
wellbeing =~						
dep_4	0.469	0.019	24.735	0.000	0.469	0.434
dep_8	0.500	0.018	27.475	0.000	0.500	0.476
dep_12	0.763	0.015	51.763	0.000	0.763	0.812
dep_16	0.829	0.016	52.273	0.000	0.829	0.818

Covariances:

	Estimate	Std.Err	z-value	P(>|z|)	Std.lv	Std.all
depression ~~						
wellbeing	0.622	0.013	47.483	0.000	0.622	0.622

Table A3. Factor Loadings for ML CFA using Hawai‘i sample, Single Factor, Method Variance Factor

Latent Variables:

	Estimate	Std.Err	z-value	P(>|z|)	Std.lv	Std.all
F1 =~						
dep_1	0.505	0.014	36.215	0.000	0.505	0.577
dep_2	0.364	0.014	26.929	0.000	0.364	0.447
dep_3	0.699	0.014	49.943	0.000	0.699	0.739
dep_4	0.276	0.019	14.771	0.000	0.276	0.255
dep_5	0.560	0.016	34.065	0.000	0.560	0.548
dep_6	0.823	0.014	57.147	0.000	0.823	0.811
dep_7	0.064	0.018	3.583	0.000	0.064	0.063
dep_8	0.211	0.018	11.551	0.000	0.211	0.201
dep_9	0.528	0.012	43.065	0.000	0.528	0.662
dep_10	0.458	0.012	36.986	0.000	0.458	0.587
dep_11	0.498	0.016	31.189	0.000	0.498	0.508
dep_12	0.501	0.015	32.940	0.000	0.501	0.533
dep_13	0.398	0.015	27.130	0.000	0.398	0.450
dep_14	0.711	0.014	49.812	0.000	0.711	0.738
dep_15	0.328	0.012	26.631	0.000	0.328	0.442
dep_16	0.516	0.017	31.217	0.000	0.516	0.509
dep_17	0.514	0.012	42.177	0.000	0.514	0.652
dep_18	0.762	0.013	57.888	0.000	0.762	0.818
dep_19	0.520	0.013	39.328	0.000	0.520	0.617
dep_20	0.590	0.014	41.676	0.000	0.590	0.646
method =~						
dep_4	0.407	0.020	19.946	0.000	0.407	0.376
dep_8	0.516	0.020	26.275	0.000	0.516	0.492
dep_12	0.551	0.016	35.262	0.000	0.551	0.586
dep_16	0.651	0.017	37.565	0.000	0.651	0.642

Covariances:

	Estimate	Std.Err	z-value	P(>|z|)	Std.lv	Std.all
F1 ~~						
method	0.000				0.000	0.000

Table A4. Factor Loadings for WLSMV CFA using Hawai‘i sample, Single Factor Model

Latent Variables:

	Estimate	Std.Err	z-value	P(>|z|)	Std.lv	Std.all
F1 =~						
dep94_1	0.608	0.019	31.452	0.000	0.608	0.608
dep94_2	0.540	0.023	23.899	0.000	0.540	0.540
dep94_3	0.816	0.012	66.042	0.000	0.816	0.816
dep94_4	0.367	0.023	15.810	0.000	0.367	0.367
dep94_5	0.612	0.017	35.454	0.000	0.612	0.612
dep94_6	0.877	0.008	105.027	0.000	0.877	0.877
dep94_7	0.011	0.027	0.399	0.690	0.011	0.011
dep94_8	0.398	0.022	17.802	0.000	0.398	0.398
dep94_9	0.793	0.016	48.986	0.000	0.793	0.793
dep94_10	0.665	0.021	32.082	0.000	0.665	0.665
dep94_11	0.603	0.019	31.035	0.000	0.603	0.603
dep94_12	0.745	0.013	56.654	0.000	0.745	0.745
dep94_13	0.526	0.021	24.616	0.000	0.526	0.526
dep94_14	0.808	0.013	63.957	0.000	0.808	0.808
dep94_15	0.618	0.023	27.285	0.000	0.618	0.618
dep94_16	0.731	0.014	51.711	0.000	0.731	0.731
dep94_17	0.727	0.019	38.858	0.000	0.727	0.727
dep94_18	0.862	0.009	94.082	0.000	0.862	0.862
dep94_19	0.764	0.015	49.311	0.000	0.764	0.764
dep94_20	0.707	0.016	43.303	0.000	0.707	0.707

Table A5. Factor Loadings for WLSMV CFA using Hawai‘i sample, Two Factor Model

Latent Variables:

	Estimate	Std.Err	z-value	P(>|z|)	Std.lv	Std.all
depression =~						
dep94_1	0.618	0.019	31.846	0.000	0.618	0.618
dep94_2	0.550	0.023	24.068	0.000	0.550	0.550
dep94_3	0.826	0.012	67.312	0.000	0.826	0.826
dep94_5	0.625	0.017	36.012	0.000	0.625	0.625
dep94_6	0.889	0.008	106.256	0.000	0.889	0.889
dep94_7	0.023	0.028	0.816	0.415	0.023	0.023
dep94_9	0.805	0.016	49.258	0.000	0.805	0.805
dep94_10	0.676	0.021	32.645	0.000	0.676	0.676
dep94_11	0.614	0.020	31.389	0.000	0.614	0.614
dep94_13	0.537	0.022	24.979	0.000	0.537	0.537
dep94_14	0.818	0.013	64.949	0.000	0.818	0.818
dep94_15	0.629	0.023	27.534	0.000	0.629	0.629
dep94_17	0.737	0.019	39.363	0.000	0.737	0.737
dep94_18	0.870	0.009	96.026	0.000	0.870	0.870
dep94_19	0.775	0.016	49.925	0.000	0.775	0.775
dep94_20	0.719	0.016	44.037	0.000	0.719	0.719
wellbeing =~						
dep94_4	0.489	0.024	20.261	0.000	0.489	0.489
dep94_8	0.525	0.023	22.937	0.000	0.525	0.525
dep94_12	0.898	0.013	67.454	0.000	0.898	0.898
dep94_16	0.876	0.014	63.462	0.000	0.876	0.876

Covariances:

	Estimate	Std.Err	z-value	P(>|z|)	Std.lv	Std.all
depression ~~						
wellbeing	0.660	0.017	38.111	0.000	0.660	0.660

Table A6. Factor Loadings for WLSMV CFA using Hawai‘i sample, Single Factor, Method Variance Factor

Latent Variables:

	Estimate	Std.Err	z-value	P(>|z|)	Std.lv	Std.all
F1 =~						
dep94_1	0.620	0.019	31.991	0.000	0.620	0.620
dep94_2	0.550	0.023	24.077	0.000	0.550	0.550
dep94_3	0.826	0.012	67.448	0.000	0.826	0.826
dep94_4	0.242	0.026	9.277	0.000	0.242	0.242
dep94_5	0.626	0.017	36.140	0.000	0.626	0.626
dep94_6	0.887	0.008	106.250	0.000	0.887	0.887
dep94_7	0.030	0.028	1.096	0.273	0.030	0.030
dep94_8	0.266	0.026	10.352	0.000	0.266	0.266
dep94_9	0.803	0.016	49.212	0.000	0.803	0.803
dep94_10	0.676	0.021	32.735	0.000	0.676	0.676
dep94_11	0.615	0.020	31.473	0.000	0.615	0.615
dep94_12	0.632	0.018	35.956	0.000	0.632	0.632
dep94_13	0.538	0.022	24.978	0.000	0.538	0.538
dep94_14	0.818	0.013	65.110	0.000	0.818	0.818
dep94_15	0.630	0.023	27.593	0.000	0.630	0.630
dep94_16	0.610	0.019	32.786	0.000	0.610	0.610
dep94_17	0.737	0.019	39.458	0.000	0.737	0.737
dep94_18	0.870	0.009	95.835	0.000	0.870	0.870
dep94_19	0.774	0.016	49.918	0.000	0.774	0.774
dep94_20	0.719	0.016	44.105	0.000	0.719	0.719
method =~						
dep94_4	0.603	0.022	27.094	0.000	0.603	0.603
dep94_8	0.632	0.021	30.053	0.000	0.632	0.632
dep94_12	0.558	0.020	28.562	0.000	0.558	0.558
dep94_16	0.593	0.019	30.912	0.000	0.593	0.593

Covariances:

	Estimate	Std.Err	z-value	P(>|z|)	Std.lv	Std.all
F1 ~~						
method	0.000				0.000	0.000

Factor Loadings for Chapter 3 Confirmatory Factor Analyses

Table A7. Factor Loadings for ML CFA using Hawai‘i sample, Trait items, Two Factor Model

Latent Variables:

	Estimate	Std.Err	z-value	P(>|z|)	Std.lv	Std.all
calmness =~						
trait1	0.648	0.012	52.172	0.000	0.648	0.782
trait6	0.573	0.014	40.130	0.000	0.573	0.644
trait7	0.660	0.013	50.447	0.000	0.660	0.764
trait10	0.678	0.013	52.825	0.000	0.678	0.788
trait13	0.679	0.013	50.561	0.000	0.679	0.765
trait16	0.534	0.015	35.189	0.000	0.534	0.580
trait19	0.571	0.013	43.267	0.000	0.571	0.683
anxiety =~						
trait2	0.379	0.014	26.674	0.000	0.379	0.457
trait3	0.501	0.013	39.761	0.000	0.501	0.639
trait4	0.658	0.016	41.435	0.000	0.658	0.660
trait5	0.477	0.014	34.128	0.000	0.477	0.565
trait8	0.595	0.014	43.455	0.000	0.595	0.684
trait9	0.618	0.015	42.560	0.000	0.618	0.674
trait11	0.527	0.015	35.471	0.000	0.527	0.583
trait12	0.506	0.014	35.136	0.000	0.506	0.579
trait14	0.183	0.017	10.621	0.000	0.183	0.191
trait15	0.535	0.012	42.977	0.000	0.535	0.679
trait17	0.591	0.014	41.951	0.000	0.591	0.666
trait18	0.568	0.015	38.414	0.000	0.568	0.622
trait20	0.687	0.015	45.593	0.000	0.687	0.709

Covariances:

	Estimate	Std.Err	z-value	P(>|z|)	Std.lv	Std.all
calmness ~~						
anxiety	0.498	0.015	33.007	0.000	0.498	0.498

Table A8. Factor Loadings for ML CFA using Hawai‘i sample, Trait items, Single Factor Model

Latent Variables:

	Estimate	Std.Err	z-value	P(>|z|)	Std.lv	Std.all
F1 =~						
trait1	0.499	0.013	37.091	0.000	0.499	0.602
trait2	0.357	0.014	25.164	0.000	0.357	0.431
trait3	0.491	0.013	38.960	0.000	0.491	0.626
trait4	0.644	0.016	40.576	0.000	0.644	0.646
trait5	0.439	0.014	31.143	0.000	0.439	0.520
trait6	0.443	0.015	29.660	0.000	0.443	0.499
trait7	0.499	0.014	35.216	0.000	0.499	0.577
trait8	0.556	0.014	40.015	0.000	0.556	0.639
trait9	0.550	0.015	36.902	0.000	0.550	0.599
trait10	0.557	0.014	40.626	0.000	0.557	0.647
trait11	0.441	0.015	28.912	0.000	0.441	0.488
trait12	0.509	0.014	35.660	0.000	0.509	0.583
trait13	0.524	0.014	36.194	0.000	0.524	0.590
trait14	0.098	0.017	5.680	0.000	0.098	0.102
trait15	0.527	0.012	42.299	0.000	0.527	0.668
trait16	0.364	0.016	22.834	0.000	0.364	0.395
trait17	0.520	0.014	35.860	0.000	0.520	0.586
trait18	0.492	0.015	32.454	0.000	0.492	0.539
trait19	0.413	0.014	29.350	0.000	0.413	0.494
trait20	0.624	0.015	40.443	0.000	0.624	0.645

Table A9. Factor Loadings for ML CFA using Hawai‘i sample, Trait items, Single Factor, Method Variance Factor

Latent Variables:

	Estimate	Std.Err	z-value	P(>|z|)	Std.lv	Std.all
F1 =~						
trait1	0.330	0.014	22.880	0.000	0.330	0.398
trait2	0.378	0.014	26.627	0.000	0.378	0.456
trait3	0.502	0.013	39.843	0.000	0.502	0.640
trait4	0.657	0.016	41.401	0.000	0.657	0.659
trait5	0.477	0.014	34.108	0.000	0.477	0.565
trait6	0.284	0.016	18.083	0.000	0.284	0.320
trait7	0.320	0.015	21.125	0.000	0.320	0.370
trait8	0.596	0.014	43.602	0.000	0.596	0.686
trait9	0.618	0.015	42.542	0.000	0.618	0.673
trait10	0.393	0.015	26.656	0.000	0.393	0.457
trait11	0.526	0.015	35.415	0.000	0.526	0.582
trait12	0.505	0.014	35.120	0.000	0.505	0.578
trait13	0.342	0.015	22.066	0.000	0.342	0.385
trait14	0.183	0.017	10.639	0.000	0.183	0.191
trait15	0.537	0.012	43.176	0.000	0.537	0.681
trait16	0.188	0.017	11.333	0.000	0.188	0.204
trait17	0.591	0.014	41.916	0.000	0.591	0.666
trait18	0.568	0.015	38.362	0.000	0.568	0.621
trait19	0.242	0.015	16.305	0.000	0.242	0.290
trait20	0.685	0.015	45.493	0.000	0.685	0.708
method =~						
trait1	0.555	0.012	45.200	0.000	0.555	0.669
trait6	0.498	0.014	34.628	0.000	0.498	0.560
trait7	0.579	0.013	44.488	0.000	0.579	0.669
trait10	0.550	0.012	44.067	0.000	0.550	0.639
trait13	0.587	0.013	44.182	0.000	0.587	0.662
trait16	0.518	0.015	33.516	0.000	0.518	0.562
trait19	0.525	0.013	39.519	0.000	0.525	0.628

Covariances:

	Estimate	Std.Err	z-value	P(>|z|)	Std.lv	Std.all
F1 ~~						
method	0.000				0.000	0.000

Table A10. Factor Loadings for ML CFA using Hawai‘i sample, State items, Two Factor Model

Latent Variables:

	Estimate	Std.Err	z-value	P(>|z|)	Std.lv	Std.all
calmness =~						
state1	0.630	0.013	48.687	0.000	0.630	0.723
state2	0.691	0.013	52.399	0.000	0.691	0.762
state5	0.721	0.015	49.442	0.000	0.721	0.732
state8	0.616	0.016	39.442	0.000	0.616	0.616
state10	0.773	0.014	56.171	0.000	0.773	0.799
state11	0.709	0.014	50.204	0.000	0.709	0.740
state15	0.820	0.014	59.369	0.000	0.820	0.829
state16	0.639	0.016	40.271	0.000	0.639	0.626
state19	0.742	0.015	48.783	0.000	0.742	0.724
state20	0.770	0.014	54.392	0.000	0.770	0.782
anxiety =~						
state3	0.610	0.014	43.206	0.000	0.610	0.678
state4	0.547	0.015	35.757	0.000	0.547	0.583
state6	0.572	0.013	44.266	0.000	0.572	0.690
state7	0.639	0.016	38.885	0.000	0.639	0.624
state9	0.352	0.017	20.446	0.000	0.352	0.356
state12	0.579	0.013	44.757	0.000	0.579	0.696
state13	0.506	0.012	43.819	0.000	0.506	0.685
state14	0.312	0.014	21.595	0.000	0.312	0.374
state17	0.696	0.015	47.494	0.000	0.696	0.727
state18	0.384	0.014	26.915	0.000	0.384	0.457

Covariances:

	Estimate	Std.Err	z-value	P(>|z|)	Std.lv	Std.all
calmness ~~						
anxiety	0.481	0.015	32.290	0.000	0.481	0.481

Table A11. Factor Loadings for ML CFA using Hawai‘i sample, State items, Single Factor Model

Latent Variables:

	Estimate	Std.Err	z-value	P(>|z|)	Std.lv	Std.all
F1 =~						
state1	0.628	0.013	48.571	0.000	0.628	0.721
state2	0.686	0.013	51.950	0.000	0.686	0.757
state3	0.422	0.015	28.642	0.000	0.422	0.469
state4	0.386	0.016	24.817	0.000	0.386	0.412
state5	0.714	0.015	48.932	0.000	0.714	0.725
state6	0.435	0.013	32.612	0.000	0.435	0.525
state7	0.450	0.017	26.622	0.000	0.450	0.440
state8	0.604	0.016	38.578	0.000	0.604	0.604
state9	-0.005	0.017	-0.286	0.775	-0.005	-0.005
state10	0.762	0.014	55.119	0.000	0.762	0.788
state11	0.695	0.014	48.889	0.000	0.695	0.725
state12	0.302	0.014	21.588	0.000	0.302	0.363
state13	0.259	0.012	20.804	0.000	0.259	0.350
state14	0.055	0.014	3.813	0.000	0.055	0.066
state15	0.820	0.014	59.565	0.000	0.820	0.830
state16	0.620	0.016	38.850	0.000	0.620	0.608
state17	0.486	0.015	31.372	0.000	0.486	0.508
state18	0.073	0.015	5.008	0.000	0.073	0.087
state19	0.719	0.015	46.872	0.000	0.719	0.703
state20	0.751	0.014	52.614	0.000	0.751	0.763

Table A12. Factor Loadings for ML CFA using Hawai‘i sample, State items, Single Factor, Method Variance Factor

Latent Variables:

	Estimate	Std.Err	z-value	P(>|z|)	Std.lv	Std.all
F1 =~						
state1	0.360	0.015	24.051	0.000	0.360	0.414
state2	0.363	0.016	23.251	0.000	0.363	0.401
state3	0.610	0.014	43.169	0.000	0.610	0.677
state4	0.540	0.015	35.268	0.000	0.540	0.577
state5	0.371	0.017	21.722	0.000	0.371	0.377
state6	0.568	0.013	43.898	0.000	0.568	0.686
state7	0.629	0.016	38.157	0.000	0.629	0.615
state8	0.265	0.018	14.950	0.000	0.265	0.265
state9	0.361	0.017	20.967	0.000	0.361	0.364
state10	0.368	0.017	21.947	0.000	0.368	0.380
state11	0.307	0.017	18.223	0.000	0.307	0.320
state12	0.584	0.013	45.342	0.000	0.584	0.703
state13	0.511	0.011	44.466	0.000	0.511	0.692
state14	0.318	0.014	22.095	0.000	0.318	0.382
state15	0.463	0.017	27.611	0.000	0.463	0.468
state16	0.242	0.018	13.322	0.000	0.242	0.237
state17	0.689	0.015	46.820	0.000	0.689	0.719
state18	0.394	0.014	27.775	0.000	0.394	0.470
state19	0.259	0.018	14.257	0.000	0.259	0.253
state20	0.315	0.017	18.257	0.000	0.315	0.320
method =~						
state1	0.514	0.013	40.319	0.000	0.514	0.590
state2	0.583	0.013	44.896	0.000	0.583	0.644
state5	0.613	0.014	42.443	0.000	0.613	0.623
state8	0.557	0.016	35.295	0.000	0.557	0.557
state10	0.679	0.014	50.186	0.000	0.679	0.703
state11	0.643	0.014	45.629	0.000	0.643	0.670
state15	0.678	0.013	51.198	0.000	0.678	0.686

state16	0.598	0.016	37.315	0.000	0.598	0.587
state19	0.716	0.015	47.153	0.000	0.716	0.699
state20	0.714	0.014	51.034	0.000	0.714	0.726

Covariances:

	Estimate	Std.Err	z-value	P(>|z|)	Std.lv	Std.all
F1 ~~						
method	0.000				0.000	0.000

Table A13. Factor Loadings for ML CFA using British sample, Trait items, Two Factor Model

Latent Variables:

	Estimate	Std.Err	z-value	P(>|z|)	Std.lv	Std.all
calmness =~						
ac1	1.012	0.030	33.246	0.000	1.012	0.700
ac6	0.949	0.031	30.641	0.000	0.949	0.658
ac7	1.088	0.028	38.435	0.000	1.088	0.777
ac10	0.773	0.023	34.092	0.000	0.773	0.713
ac13	1.096	0.030	37.075	0.000	1.096	0.758
ac16	1.065	0.029	36.748	0.000	1.065	0.753
ac19	1.168	0.032	35.989	0.000	1.168	0.742
anxiety =~						
ac2	0.589	0.026	22.633	0.000	0.589	0.505
ac3	0.959	0.027	35.604	0.000	0.959	0.727
ac4	1.164	0.036	32.224	0.000	1.164	0.674
ac5	1.297	0.033	39.297	0.000	1.297	0.779
ac8	1.226	0.028	43.494	0.000	1.226	0.833
ac9	1.029	0.026	38.947	0.000	1.029	0.774
ac11	1.091	0.029	38.270	0.000	1.091	0.765
ac12	1.069	0.028	37.788	0.000	1.069	0.758
ac14	1.088	0.031	35.384	0.000	1.088	0.723
ac15	0.865	0.024	36.214	0.000	0.865	0.736
ac17	1.146	0.026	43.456	0.000	1.146	0.833
ac18	1.291	0.030	43.772	0.000	1.291	0.837
ac20	1.190	0.027	43.916	0.000	1.190	0.838

Covariances:

	Estimate	Std.Err	z-value	P(>|z|)	Std.lv	Std.all
calmness ~~						
anxiety	0.877	0.008	114.179	0.000	0.877	0.877

Table A14. Factor Loadings for ML CFA using British sample, Trait items, Single Factor Model

Latent Variables:

	Estimate	Std.Err	z-value	P(>|z|)	Std.lv	Std.all
F1 =~						
ac1	0.914	0.031	29.733	0.000	0.914	0.632
ac2	0.595	0.026	22.947	0.000	0.595	0.510
ac3	0.955	0.027	35.479	0.000	0.955	0.724
ac4	1.151	0.036	31.834	0.000	1.151	0.667
ac5	1.286	0.033	38.905	0.000	1.286	0.773
ac6	0.934	0.030	30.660	0.000	0.934	0.648
ac7	1.040	0.028	36.756	0.000	1.040	0.742
ac8	1.222	0.028	43.317	0.000	1.222	0.830
ac9	1.013	0.027	38.165	0.000	1.013	0.762
ac10	0.694	0.023	30.224	0.000	0.694	0.641
ac11	1.074	0.029	37.482	0.000	1.074	0.753
ac12	1.069	0.028	37.880	0.000	1.069	0.758
ac13	1.007	0.030	33.671	0.000	1.007	0.696
ac14	1.077	0.031	34.965	0.000	1.077	0.716
ac15	0.875	0.024	36.858	0.000	0.875	0.744
ac16	0.966	0.029	32.845	0.000	0.966	0.683
ac17	1.136	0.026	42.964	0.000	1.136	0.826
ac18	1.276	0.030	43.062	0.000	1.276	0.827
ac19	1.076	0.033	32.867	0.000	1.076	0.684
ac20	1.171	0.027	42.937	0.000	1.171	0.825

Table A15. Factor Loadings for ML CFA using British sample, Trait items, Single Factor, Method Variance Factor

Latent Variables:

	Estimate	Std.Err	z-value	P(>|z|)	Std.lv	Std.all
F1 =~						
ac1	0.857	0.031	27.362	0.000	0.857	0.593
ac2	0.591	0.026	22.745	0.000	0.591	0.507
ac3	0.961	0.027	35.701	0.000	0.961	0.728
ac4	1.164	0.036	32.239	0.000	1.164	0.675
ac5	1.300	0.033	39.419	0.000	1.300	0.781
ac6	0.912	0.031	29.626	0.000	0.912	0.632
ac7	1.004	0.029	34.921	0.000	1.004	0.717
ac8	1.228	0.028	43.569	0.000	1.228	0.834
ac9	1.027	0.026	38.879	0.000	1.027	0.773
ac10	0.646	0.023	27.541	0.000	0.646	0.596
ac11	1.090	0.029	38.228	0.000	1.090	0.764
ac12	1.068	0.028	37.763	0.000	1.068	0.758
ac13	0.948	0.031	31.014	0.000	0.948	0.655
ac14	1.088	0.031	35.386	0.000	1.088	0.723
ac15	0.863	0.024	36.091	0.000	0.863	0.734
ac16	0.905	0.030	30.087	0.000	0.905	0.640
ac17	1.146	0.026	43.463	0.000	1.146	0.833
ac18	1.291	0.029	43.766	0.000	1.291	0.837
ac19	1.019	0.033	30.544	0.000	1.019	0.648
ac20	1.187	0.027	43.800	0.000	1.187	0.837
method =~						
ac1	0.536	0.033	16.026	0.000	0.536	0.371
ac6	0.220	0.033	6.583	0.000	0.220	0.152
ac7	0.353	0.029	12.276	0.000	0.353	0.252
ac10	0.469	0.025	18.911	0.000	0.469	0.433
ac13	0.604	0.031	19.395	0.000	0.604	0.418
ac16	0.625	0.031	20.207	0.000	0.625	0.442
ac19	0.566	0.035	16.392	0.000	0.566	0.360

Covariances:

	Estimate	Std.Err	z-value	P(>|z|)	Std.lv	Std.all
F1 ~~						
method	0.000				0.000	0.000

Table A16. Factor Loadings for ML CFA using British sample, State items, Two Factor Model

Latent Variables:

	Estimate	Std.Err	z-value	P(>|z|)	Std.lv	Std.all
calmness =~						
cseq1	1.133	0.027	41.793	0.000	1.133	0.882
cseq2	1.163	0.030	38.483	0.000	1.163	0.838
cseq5	1.158	0.031	37.428	0.000	1.158	0.823
cseq8	1.105	0.031	35.364	0.000	1.105	0.793
cseq10	1.119	0.031	36.089	0.000	1.119	0.803
cseq11	1.126	0.033	34.582	0.000	1.126	0.781
cseq15	1.185	0.030	39.551	0.000	1.185	0.852
cseq16	1.218	0.032	38.188	0.000	1.218	0.834
cseq19	1.154	0.031	37.692	0.000	1.154	0.827
cseq20	1.179	0.029	40.397	0.000	1.179	0.864
anxiety =~						
cseq3	1.181	0.032	36.848	0.000	1.181	0.817
cseq4	1.132	0.032	35.940	0.000	1.132	0.803
cseq6	1.207	0.044	27.182	0.000	1.207	0.656
cseq7	1.195	0.033	36.264	0.000	1.195	0.808
cseq9	1.183	0.034	34.500	0.000	1.183	0.781
cseq12	1.112	0.033	34.024	0.000	1.112	0.774
cseq13	1.290	0.032	40.120	0.000	1.290	0.862
cseq14	1.293	0.038	34.226	0.000	1.293	0.777
cseq17	1.230	0.037	33.638	0.000	1.230	0.768
cseq18	1.212	0.035	34.295	0.000	1.212	0.778

Covariances:

	Estimate	Std.Err	z-value	P(>|z|)	Std.lv	Std.all
calmness ~~						
anxiety	0.900	0.007	136.703	0.000	0.900	0.900

Table A17. Factor Loadings for ML CFA using British sample, State items, Single Factor Model

Latent Variables:

	Estimate	Std.Err	z-value	P(>|z|)	Std.lv	Std.all
F1 =~						
cseq1	1.134	0.027	42.006	0.000	1.134	0.883
cseq2	1.162	0.030	38.541	0.000	1.162	0.837
cseq3	1.174	0.032	36.760	0.000	1.174	0.812
cseq4	1.138	0.031	36.468	0.000	1.138	0.807
cseq5	1.141	0.031	36.723	0.000	1.141	0.811
cseq6	1.201	0.044	27.202	0.000	1.201	0.653
cseq7	1.155	0.033	34.746	0.000	1.155	0.781
cseq8	1.079	0.032	34.241	0.000	1.079	0.774
cseq9	1.128	0.035	32.440	0.000	1.128	0.745
cseq10	1.095	0.031	35.064	0.000	1.095	0.786
cseq11	1.112	0.033	34.063	0.000	1.112	0.771
cseq12	1.033	0.033	30.892	0.000	1.033	0.719
cseq13	1.216	0.033	36.797	0.000	1.216	0.812
cseq14	1.228	0.038	32.043	0.000	1.228	0.738
cseq15	1.169	0.030	38.814	0.000	1.169	0.841
cseq16	1.198	0.032	37.327	0.000	1.198	0.820
cseq17	1.174	0.037	31.728	0.000	1.174	0.733
cseq18	1.140	0.036	31.638	0.000	1.140	0.731
cseq19	1.111	0.031	35.673	0.000	1.111	0.796
cseq20	1.132	0.030	37.970	0.000	1.132	0.829

Table A18. Factor Loadings for ML CFA using British sample, State items, Single Factor, Method Variance Factor

Latent Variables:

	Estimate	Std.Err	z-value	P(>|z|)	Std.lv	Std.all
F1 =~						
cseq1	1.091	0.028	39.181	0.000	1.091	0.849
cseq2	1.111	0.031	35.736	0.000	1.111	0.800
cseq3	1.184	0.032	37.059	0.000	1.184	0.819
cseq4	1.133	0.031	36.008	0.000	1.133	0.803
cseq5	1.064	0.032	32.890	0.000	1.064	0.756
cseq6	1.209	0.044	27.278	0.000	1.209	0.657
cseq7	1.192	0.033	36.206	0.000	1.192	0.806
cseq8	0.988	0.033	30.038	0.000	0.988	0.708
cseq9	1.184	0.034	34.559	0.000	1.184	0.782
cseq10	1.002	0.033	30.656	0.000	1.002	0.719
cseq11	1.040	0.034	30.745	0.000	1.040	0.720
cseq12	1.110	0.033	33.986	0.000	1.110	0.773
cseq13	1.280	0.032	39.683	0.000	1.280	0.855
cseq14	1.282	0.038	33.848	0.000	1.282	0.770
cseq15	1.093	0.031	34.822	0.000	1.093	0.786
cseq16	1.109	0.034	33.083	0.000	1.109	0.759
cseq17	1.221	0.037	33.320	0.000	1.221	0.762
cseq18	1.209	0.035	34.217	0.000	1.209	0.776
cseq19	0.968	0.033	29.154	0.000	0.968	0.694
cseq20	0.986	0.032	30.809	0.000	0.986	0.723
method =~						
cseq1	0.284	0.021	13.236	0.000	0.284	0.221
cseq2	0.318	0.026	12.235	0.000	0.318	0.229
cseq5	0.427	0.028	15.405	0.000	0.427	0.304
cseq8	0.477	0.029	16.314	0.000	0.477	0.342
cseq10	0.498	0.029	17.468	0.000	0.498	0.358
cseq11	0.407	0.030	13.384	0.000	0.407	0.282
cseq15	0.437	0.026	17.000	0.000	0.437	0.315

cseq16	0.489	0.028	17.337	0.000	0.489	0.335
cseq19	0.746	0.027	27.519	0.000	0.746	0.535
cseq20	0.776	0.025	31.506	0.000	0.776	0.569

Covariances:

	Estimate	Std.Err	z-value	P(>|z|)	Std.lv	Std.all
F1 ~~						
method	0.000				0.000	0.000

Table A19. Factor Loadings for WLSMV CFA using Hawai‘i sample, Trait items, Two Factor Model

Latent Variables:

	Estimate	Std.Err	z-value	P(>|z|)	Std.lv	Std.all
calmness =~						
trait1	0.834	0.007	126.533	0.000	0.834	0.834
trait6	0.696	0.010	67.801	0.000	0.696	0.696
trait7	0.809	0.008	106.417	0.000	0.809	0.809
trait10	0.871	0.006	136.999	0.000	0.871	0.871
trait13	0.815	0.008	108.704	0.000	0.815	0.815
trait16	0.596	0.012	50.925	0.000	0.596	0.596
trait19	0.721	0.010	73.735	0.000	0.721	0.721
anxiety =~						
trait2	0.495	0.015	33.295	0.000	0.495	0.495
trait3	0.738	0.011	67.096	0.000	0.738	0.738
trait4	0.726	0.011	69.165	0.000	0.726	0.726
trait5	0.625	0.013	48.476	0.000	0.625	0.625
trait8	0.733	0.010	74.269	0.000	0.733	0.733
trait9	0.714	0.010	70.054	0.000	0.714	0.714
trait11	0.604	0.013	47.013	0.000	0.604	0.604
trait12	0.671	0.012	55.769	0.000	0.671	0.671
trait14	0.161	0.018	9.007	0.000	0.161	0.161
trait15	0.782	0.010	77.400	0.000	0.782	0.782
trait17	0.710	0.010	70.402	0.000	0.710	0.710
trait18	0.658	0.012	54.629	0.000	0.658	0.658
trait20	0.755	0.009	80.086	0.000	0.755	0.755

Covariances:

	Estimate	Std.Err	z-value	P(>|z|)	Std.lv	Std.all
calmness ~~						
anxiety	0.514	0.014	35.670	0.000	0.514	0.514

Table A20. Factor Loadings for WLSMV CFA using Hawai‘i sample, Trait items, Single Factor Model

Latent Variables:

	Estimate	Std.Err	z-value	P(>|z|)	Std.lv	Std.all
F1 =~						
trait1	0.760	0.007	102.281	0.000	0.760	0.760
trait2	0.445	0.014	30.906	0.000	0.445	0.445
trait3	0.681	0.011	61.896	0.000	0.681	0.681
trait4	0.667	0.010	63.562	0.000	0.667	0.667
trait5	0.564	0.013	43.930	0.000	0.564	0.564
trait6	0.604	0.011	56.137	0.000	0.604	0.604
trait7	0.732	0.008	88.289	0.000	0.732	0.732
trait8	0.670	0.010	66.302	0.000	0.670	0.670
trait9	0.646	0.011	61.212	0.000	0.646	0.646
trait10	0.799	0.007	114.673	0.000	0.799	0.799
trait11	0.527	0.013	40.704	0.000	0.527	0.527
trait12	0.616	0.012	53.021	0.000	0.616	0.616
trait13	0.733	0.008	88.926	0.000	0.733	0.733
trait14	0.099	0.017	5.963	0.000	0.099	0.099
trait15	0.718	0.010	70.822	0.000	0.718	0.718
trait16	0.501	0.012	41.658	0.000	0.501	0.501
trait17	0.645	0.010	61.532	0.000	0.645	0.645
trait18	0.592	0.012	48.239	0.000	0.592	0.592
trait19	0.628	0.010	60.735	0.000	0.628	0.628
trait20	0.691	0.010	70.588	0.000	0.691	0.691

Table A21. Factor Loadings for WLSMV CFA using Hawai‘i sample, Trait items, Single Factor, Method Variance Factor

Latent Variables:

	Estimate	Std.Err	z-value	P(>|z|)	Std.lv	Std.all
F1 =~						
trait1	0.447	0.015	29.719	0.000	0.447	0.447
trait2	0.495	0.015	33.278	0.000	0.495	0.495
trait3	0.738	0.011	67.135	0.000	0.738	0.738
trait4	0.726	0.010	69.162	0.000	0.726	0.726
trait5	0.625	0.013	48.510	0.000	0.625	0.625
trait6	0.352	0.016	21.594	0.000	0.352	0.352
trait7	0.415	0.015	27.044	0.000	0.415	0.415
trait8	0.734	0.010	74.530	0.000	0.734	0.734
trait9	0.714	0.010	70.179	0.000	0.714	0.714
trait10	0.513	0.015	35.154	0.000	0.513	0.513
trait11	0.605	0.013	47.086	0.000	0.605	0.605
trait12	0.671	0.012	55.780	0.000	0.671	0.671
trait13	0.431	0.015	28.504	0.000	0.431	0.431
trait14	0.163	0.018	9.097	0.000	0.163	0.163
trait15	0.782	0.010	77.580	0.000	0.782	0.782
trait16	0.213	0.017	12.589	0.000	0.213	0.213
trait17	0.710	0.010	70.505	0.000	0.710	0.710
trait18	0.659	0.012	54.679	0.000	0.659	0.659
trait19	0.330	0.016	20.282	0.000	0.330	0.330
trait20	0.755	0.009	80.076	0.000	0.755	0.755
method =~						
trait1	0.700	0.010	66.889	0.000	0.700	0.700
trait6	0.603	0.012	50.687	0.000	0.603	0.603
trait7	0.696	0.010	66.834	0.000	0.696	0.696
trait10	0.680	0.011	59.710	0.000	0.680	0.680
trait13	0.687	0.010	66.449	0.000	0.687	0.687
trait16	0.614	0.011	53.573	0.000	0.614	0.614
trait19	0.663	0.011	59.780	0.000	0.663	0.663

Covariances:

	Estimate	Std.Err	z-value	P(>|z|)	Std.lv	Std.all
F1 ~~						
method	0.000				0.000	0.000

Table A22. Factor Loadings for WLSMV CFA using Hawai‘i sample, State items, Two Factor Model

Latent Variables:

	Estimate	Std.Err	z-value	P(>|z|)	Std.lv	Std.all
calmness =~						
state1	0.808	0.007	120.694	0.000	0.808	0.808
state2	0.836	0.006	136.515	0.000	0.836	0.836
state5	0.780	0.008	102.044	0.000	0.780	0.780
state8	0.656	0.010	62.761	0.000	0.656	0.656
state10	0.837	0.006	134.864	0.000	0.837	0.837
state11	0.776	0.007	104.826	0.000	0.776	0.776
state15	0.875	0.005	174.635	0.000	0.875	0.875
state16	0.661	0.010	65.892	0.000	0.661	0.661
state19	0.812	0.007	124.603	0.000	0.812	0.812
state20	0.863	0.005	161.321	0.000	0.863	0.863
anxiety =~						
state3	0.772	0.011	70.974	0.000	0.772	0.772
state4	0.690	0.014	51.063	0.000	0.690	0.690
state6	0.853	0.010	82.768	0.000	0.853	0.853
state7	0.716	0.012	59.946	0.000	0.716	0.716
state9	0.300	0.018	16.433	0.000	0.300	0.300
state12	0.770	0.011	69.465	0.000	0.770	0.770
state13	0.806	0.012	66.633	0.000	0.806	0.806
state14	0.382	0.020	19.395	0.000	0.382	0.382
state17	0.827	0.009	88.453	0.000	0.827	0.827
state18	0.472	0.018	26.256	0.000	0.472	0.472

Covariances:

	Estimate	Std.Err	z-value	P(>|z|)	Std.lv	Std.all
calmness ~~						
anxiety	0.508	0.014	36.296	0.000	0.508	0.508

Table A23. Factor Loadings for WLSMV CFA using Hawai‘i sample, State items, Single Factor Model

Latent Variables:

	Estimate	Std.Err	z-value	P(>|z|)	Std.lv	Std.all
F1 =~						
state1	0.786	0.007	114.891	0.000	0.786	0.786
state2	0.814	0.006	128.856	0.000	0.814	0.814
state3	0.618	0.012	53.181	0.000	0.618	0.618
state4	0.554	0.014	40.565	0.000	0.554	0.554
state5	0.758	0.008	97.307	0.000	0.758	0.758
state6	0.703	0.011	63.772	0.000	0.703	0.703
state7	0.577	0.013	46.110	0.000	0.577	0.577
state8	0.626	0.010	59.764	0.000	0.626	0.626
state9	0.125	0.017	7.370	0.000	0.125	0.125
state10	0.819	0.006	128.200	0.000	0.819	0.819
state11	0.750	0.008	99.132	0.000	0.750	0.750
state12	0.617	0.013	48.748	0.000	0.617	0.617
state13	0.642	0.014	46.319	0.000	0.642	0.642
state14	0.215	0.018	11.743	0.000	0.215	0.215
state15	0.856	0.005	167.242	0.000	0.856	0.856
state16	0.633	0.010	62.375	0.000	0.633	0.633
state17	0.683	0.011	64.215	0.000	0.683	0.683
state18	0.281	0.018	15.863	0.000	0.281	0.281
state19	0.790	0.007	118.620	0.000	0.790	0.790
state20	0.845	0.006	153.336	0.000	0.845	0.845

Table A24. Factor Loadings for WLSMV CFA using Hawai‘i sample, State items, Single Factor, Method Variance Factor

Latent Variables:

	Estimate	Std.Err	z-value	P(>|z|)	Std.lv	Std.all
F1 =~						
state1	0.494	0.015	33.438	0.000	0.494	0.494
state2	0.480	0.015	32.154	0.000	0.480	0.480
state3	0.769	0.011	71.140	0.000	0.769	0.769
state4	0.687	0.013	50.948	0.000	0.687	0.687
state5	0.448	0.015	29.224	0.000	0.448	0.448
state6	0.848	0.010	82.383	0.000	0.848	0.848
state7	0.712	0.012	59.606	0.000	0.712	0.712
state8	0.316	0.017	18.748	0.000	0.316	0.316
state9	0.307	0.018	16.885	0.000	0.307	0.307
state10	0.446	0.015	29.779	0.000	0.446	0.446
state11	0.375	0.016	23.295	0.000	0.375	0.375
state12	0.770	0.011	69.872	0.000	0.770	0.770
state13	0.806	0.012	67.207	0.000	0.806	0.806
state14	0.388	0.020	19.784	0.000	0.388	0.388
state15	0.535	0.014	39.549	0.000	0.535	0.535
state16	0.282	0.018	15.828	0.000	0.282	0.282
state17	0.824	0.009	87.925	0.000	0.824	0.824
state18	0.481	0.018	26.981	0.000	0.481	0.481
state19	0.262	0.017	15.319	0.000	0.262	0.262
state20	0.331	0.016	20.290	0.000	0.331	0.331
method =~						
state1	0.632	0.011	58.811	0.000	0.632	0.632
state2	0.680	0.010	66.715	0.000	0.680	0.680
state5	0.633	0.011	60.296	0.000	0.633	0.633
state8	0.580	0.012	48.616	0.000	0.580	0.580
state10	0.708	0.009	77.160	0.000	0.708	0.708
state11	0.684	0.010	69.474	0.000	0.684	0.684
state15	0.688	0.010	71.469	0.000	0.688	0.688

state16	0.614	0.012	52.912	0.000	0.614	0.614
state19	0.802	0.008	102.720	0.000	0.802	0.802
state20	0.814	0.008	107.205	0.000	0.814	0.814

Covariances:

	Estimate	Std.Err	z-value	P(>|z|)	Std.lv	Std.all
F1 ~~						
method	0.000			0.000	0.000	

Table A25. Factor Loadings for WLSMV CFA using British sample, Trait items, Two Factor Model

Latent Variables:

	Estimate	Std.Err	z-value	P(>|z|)	Std.lv	Std.all
calmness =~						
ac1	0.681	0.016	41.541	0.000	0.681	0.681
ac6	0.546	0.019	28.413	0.000	0.546	0.546
ac7	0.723	0.014	49.925	0.000	0.723	0.723
ac10	0.780	0.015	53.356	0.000	0.780	0.780
ac13	0.766	0.015	51.619	0.000	0.766	0.766
ac16	0.797	0.014	56.012	0.000	0.797	0.797
ac19	0.727	0.016	46.416	0.000	0.727	0.727
anxiety =~						
ac2	0.436	0.020	21.433	0.000	0.436	0.436
ac3	0.623	0.016	38.291	0.000	0.623	0.623
ac4	0.673	0.017	39.462	0.000	0.673	0.673
ac5	0.698	0.016	43.054	0.000	0.698	0.698
ac8	0.770	0.012	62.698	0.000	0.770	0.770
ac9	0.716	0.013	55.691	0.000	0.716	0.716
ac11	0.661	0.014	45.771	0.000	0.661	0.661
ac12	0.629	0.014	44.708	0.000	0.629	0.629
ac14	0.452	0.017	26.281	0.000	0.452	0.452
ac15	0.705	0.015	46.510	0.000	0.705	0.705
ac17	0.717	0.012	58.400	0.000	0.717	0.717
ac18	0.769	0.012	63.290	0.000	0.769	0.769
ac20	0.756	0.012	63.099	0.000	0.756	0.756

Covariances:

	Estimate	Std.Err	z-value	P(>|z|)	Std.lv	Std.all
calmness ~~						
anxiety	0.736	0.012	61.265	0.000	0.736	0.736

Table A26. Factor Loadings for WLSMV CFA using British sample, Trait items, Single Factor Model

Latent Variables:

	Estimate	Std.Err	z-value	P(>|z|)	Std.lv	Std.all
F1 =~						
ac1	0.613	0.017	37.056	0.000	0.613	0.613
ac2	0.423	0.020	21.475	0.000	0.423	0.423
ac3	0.604	0.016	37.823	0.000	0.604	0.604
ac4	0.647	0.017	39.105	0.000	0.647	0.647
ac5	0.668	0.016	41.881	0.000	0.668	0.668
ac6	0.484	0.018	26.836	0.000	0.484	0.484
ac7	0.650	0.014	46.314	0.000	0.650	0.650
ac8	0.742	0.012	61.523	0.000	0.742	0.742
ac9	0.689	0.013	53.729	0.000	0.689	0.689
ac10	0.706	0.015	47.472	0.000	0.706	0.706
ac11	0.634	0.014	44.184	0.000	0.634	0.634
ac12	0.604	0.014	43.861	0.000	0.604	0.604
ac13	0.695	0.015	46.704	0.000	0.695	0.695
ac14	0.432	0.017	25.815	0.000	0.432	0.432
ac15	0.682	0.015	46.467	0.000	0.682	0.682
ac16	0.724	0.014	50.417	0.000	0.724	0.724
ac17	0.690	0.012	56.667	0.000	0.690	0.690
ac18	0.740	0.012	61.106	0.000	0.740	0.740
ac19	0.657	0.016	42.151	0.000	0.657	0.657
ac20	0.728	0.012	60.852	0.000	0.728	0.728

Table A27. Factor Loadings for WLSMV CFA using British sample, Trait items, Single Factor, Method Variance Factor

Latent Variables:

	Estimate	Std.Err	z-value	P(>|z|)	Std.lv	Std.all
F1 =~						
ac1	0.470	0.020	23.030	0.000	0.470	0.470
ac2	0.436	0.020	21.415	0.000	0.436	0.436
ac3	0.623	0.016	38.301	0.000	0.623	0.623
ac4	0.673	0.017	39.419	0.000	0.673	0.673
ac5	0.698	0.016	43.078	0.000	0.698	0.698
ac6	0.413	0.019	21.322	0.000	0.413	0.413
ac7	0.576	0.016	35.371	0.000	0.576	0.576
ac8	0.770	0.012	62.720	0.000	0.770	0.770
ac9	0.716	0.013	55.758	0.000	0.716	0.716
ac10	0.542	0.019	28.743	0.000	0.542	0.542
ac11	0.661	0.014	45.794	0.000	0.661	0.661
ac12	0.629	0.014	44.800	0.000	0.629	0.629
ac13	0.581	0.018	31.861	0.000	0.581	0.581
ac14	0.452	0.017	26.298	0.000	0.452	0.452
ac15	0.705	0.015	46.429	0.000	0.705	0.705
ac16	0.576	0.018	31.674	0.000	0.576	0.576
ac17	0.717	0.012	58.479	0.000	0.717	0.717
ac18	0.769	0.012	63.349	0.000	0.769	0.769
ac19	0.533	0.019	28.054	0.000	0.533	0.533
ac20	0.756	0.012	63.130	0.000	0.756	0.756
method =~						
ac1	0.542	0.018	30.200	0.000	0.542	0.542
ac6	0.339	0.022	15.631	0.000	0.339	0.339
ac7	0.360	0.019	18.647	0.000	0.360	0.360
ac10	0.614	0.018	34.372	0.000	0.614	0.614
ac13	0.471	0.018	26.202	0.000	0.471	0.471
ac16	0.564	0.016	34.387	0.000	0.564	0.564
ac19	0.499	0.019	26.878	0.000	0.499	0.499

Covariances:

	Estimate	Std.Err	z-value	P(>|z|)	Std.lv	Std.all
F1 ~~						
method	0.000				0.000	0.000

Table A28. Factor Loadings for WLSMV CFA using British sample, State items, Two Factor Model

Latent Variables:

	Estimate	Std.Err	z-value	P(>|z|)	Std.lv	Std.all
calmness =~						
cseq1	0.807	0.011	76.366	0.000	0.807	0.807
cseq2	0.754	0.014	55.223	0.000	0.754	0.754
cseq5	0.775	0.013	59.425	0.000	0.775	0.775
cseq8	0.651	0.014	45.950	0.000	0.651	0.651
cseq10	0.691	0.013	52.175	0.000	0.691	0.691
cseq11	0.678	0.015	44.564	0.000	0.678	0.678
cseq15	0.832	0.011	75.956	0.000	0.832	0.832
cseq16	0.817	0.013	65.286	0.000	0.817	0.817
cseq19	0.708	0.011	61.985	0.000	0.708	0.708
cseq20	0.817	0.010	82.212	0.000	0.817	0.817
anxiety =~						
cseq3	0.823	0.014	56.804	0.000	0.823	0.823
cseq4	0.716	0.019	37.667	0.000	0.716	0.716
cseq6	0.754	0.018	41.668	0.000	0.754	0.754
cseq7	0.713	0.015	47.757	0.000	0.713	0.713
cseq9	0.703	0.014	49.235	0.000	0.703	0.703
cseq12	0.709	0.013	54.284	0.000	0.709	0.709
cseq13	0.830	0.012	70.723	0.000	0.830	0.830
cseq14	0.789	0.015	51.659	0.000	0.789	0.789
cseq17	0.793	0.013	59.193	0.000	0.793	0.793
cseq18	0.587	0.018	33.396	0.000	0.587	0.587

Covariances:

	Estimate	Std.Err	z-value	P(>|z|)	Std.lv	Std.all
calmness ~~						
anxiety	0.725	0.011	64.613	0.000	0.725	0.725

Table A29. Factor Loadings for WLSMV CFA using British sample, State items, Single Factor Model

Latent Variables:

	Estimate	Std.Err	z-value	P(>|z|)	Std.lv	Std.all
F1 =~						
cseq1	0.766	0.010	72.970	0.000	0.766	0.766
cseq2	0.716	0.013	53.554	0.000	0.716	0.716
cseq3	0.754	0.014	54.670	0.000	0.754	0.754
cseq4	0.653	0.017	37.843	0.000	0.653	0.653
cseq5	0.740	0.013	57.976	0.000	0.740	0.740
cseq6	0.691	0.017	40.803	0.000	0.691	0.691
cseq7	0.656	0.015	44.579	0.000	0.656	0.656
cseq8	0.612	0.014	43.980	0.000	0.612	0.612
cseq9	0.640	0.014	45.181	0.000	0.640	0.640
cseq10	0.650	0.013	49.290	0.000	0.650	0.650
cseq11	0.643	0.015	43.242	0.000	0.643	0.643
cseq12	0.652	0.013	48.586	0.000	0.652	0.652
cseq13	0.763	0.012	62.576	0.000	0.763	0.763
cseq14	0.722	0.015	48.010	0.000	0.722	0.722
cseq15	0.792	0.011	72.870	0.000	0.792	0.792
cseq16	0.782	0.012	63.134	0.000	0.782	0.782
cseq17	0.733	0.013	55.014	0.000	0.733	0.733
cseq18	0.517	0.017	30.626	0.000	0.517	0.517
cseq19	0.665	0.012	57.822	0.000	0.665	0.665
cseq20	0.777	0.010	75.734	0.000	0.777	0.777

Table A30. Factor Loadings for WLSMV CFA using British sample, State items, Single Factor, Method Variance Factor

Latent Variables:

	Estimate	Std.Err	z-value	P(>|z|)	Std.lv	Std.all
F1 =~						
cseq1	0.668	0.014	48.545	0.000	0.668	0.668
cseq2	0.608	0.017	36.321	0.000	0.608	0.608
cseq3	0.816	0.014	56.914	0.000	0.816	0.816
cseq4	0.710	0.019	37.485	0.000	0.710	0.710
cseq5	0.628	0.016	40.425	0.000	0.628	0.628
cseq6	0.750	0.018	41.778	0.000	0.750	0.750
cseq7	0.710	0.015	47.689	0.000	0.710	0.710
cseq8	0.455	0.017	26.630	0.000	0.455	0.455
cseq9	0.702	0.014	49.390	0.000	0.702	0.702
cseq10	0.486	0.017	28.372	0.000	0.486	0.486
cseq11	0.529	0.018	29.891	0.000	0.529	0.529
cseq12	0.710	0.013	54.484	0.000	0.710	0.710
cseq13	0.829	0.012	70.847	0.000	0.829	0.829
cseq14	0.787	0.015	51.904	0.000	0.787	0.787
cseq15	0.655	0.015	44.842	0.000	0.655	0.655
cseq16	0.609	0.016	37.349	0.000	0.609	0.609
cseq17	0.790	0.013	59.151	0.000	0.790	0.790
cseq18	0.590	0.018	33.694	0.000	0.590	0.590
cseq19	0.368	0.017	21.217	0.000	0.368	0.368
cseq20	0.479	0.017	28.030	0.000	0.479	0.479
method =~						
cseq1	0.397	0.018	21.964	0.000	0.397	0.397
cseq2	0.404	0.021	19.454	0.000	0.404	0.404
cseq5	0.417	0.018	22.928	0.000	0.417	0.417
cseq8	0.487	0.019	25.357	0.000	0.487	0.487
cseq10	0.512	0.018	27.906	0.000	0.512	0.512
cseq11	0.400	0.020	20.451	0.000	0.400	0.400
cseq15	0.480	0.017	28.528	0.000	0.480	0.480

cseq16	0.539	0.016	32.944	0.000	0.539	0.539
cseq19	0.726	0.014	53.118	0.000	0.726	0.726
cseq20	0.747	0.013	59.084	0.000	0.747	0.747

Covariances:

	Estimate	Std.Err	z-value	P(>|z|)	Std.lv	Std.all
F1 ~~						
method	0.000			0.000	0.000	