UNIVERSITY OF CHICHESTER

An accredited institution of the
UNIVERSITY OF SOUTHAMPTON

DEPARTMENT OF SPORT AND EXERCISE SCIENCES

THE DEVELOPMENT OF AN INSTRUMENT TO MEASURE MENTAL TOUGHNESS IN SPORT

by
Phillip D. J. Birch

Thesis for the Doctor of Philosophy

This thesis has been completed as a requirement for a higher degree of the University of Southampton.

January 2014

UNIVERSITY OF CHICHESTER

ABSTRACT

DEPARTMENT OF SPORT AND EXERCISE SCIENCES

Sport and Exercise Psychology

Doctor of Philosophy

THE DEVELOPMENT OF AN INSTRUMENT TO MEASURE MENTAL TOUGHNESS IN SPORT

by Phillip D. J. Birch

ABSTRACT

The Mental Toughness Questionnaire-48 (MTQ48; Earle, 2006) is the most utilised instrument to measure Mental Toughness (MT) in sport (Gucciardi, Hanton, & Mallett, 2012). To date, preliminary research examining the factorial validity of the MTQ48 has yielded equivocal findings (Gucciardi *et al*; Horsburgh, Schermer, Veselka, & Vernon, 2009) regarding its ability to measure the 4/6C's model of MT (Clough, Earle, & Sewell, 2002; Earle). The primary aim of this thesis was to provide a truly comprehensive examination of the factorial validity of the MTQ48 in an effort to provide an adequate measure of the 4/6C's model of MT. Study 1 re-examined the factor structure of the MTQ48 by using a very large sample of competitive student athletes and found little support for its factorial validity. Inspection of item content revealed concerns regarding the adequacy of MTQ48 items to represent the 4/6C's model of MT. Study 2 developed the University of Chichester Mental Toughness Questionnaire (UCMTQ) whereby items were generated to better represent the 4/6C's model of MT. Although the UCMTQ's factorial validity was superior to the tested models of the MTQ48, the results of Study 2 provided little support for its factorial validity. It was concluded that the poor factorial validity of the UCMTQ could have been due to the inadequacy of the factor definitions developed by Clough, Marchant, and Earle (2007) to represent the core traits underpinning the 4/6C's model of MT (challenge, commitment, control, and confidence).

Study 3 developed the Hardiness Confidence Questionnaire (HCQ) whereby factor definitions and items were regenerated to better represent the core traits of challenge, commitment, control, and confidence. Although little support was found for the hypothesised models and structural models of the HCQ, results revealed preliminary support for the factorial validity of the first-order four factor revised model when using more liberal model fit thresholds. Given that HCQ factor loadings were generally strong, it was concluded that acceptable factorial validity had been established. Based upon the findings of the first-order four factor revised model of the HCQ, Study 4 constructed the Revised Hardiness Confidence Questionnaire (HCQ-R). Findings revealed preliminary support for the HCQ-R's convergent validity when using more liberal criteria and demonstrated the instrument's ability to predict pre-competitive challenge states and pre-competitive self-confidence and somatic anxiety. However, findings revealed little support for the HCQ-R's test-retest reliability and its ability to predict coping styles in sport. Although this thesis provides the first examination of the core traits thought to underpin MT in sport, the equivocal construct validity findings suggest that further examination of the revised 4C's model of MT is needed before it can be considered a valid conceptualisation of MT in sport.

CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF APPENDICES

DECLARATION OF AUTHORSHIP

I, PHILLIP DAVID JOHN BIRCH

declare that the thesis entitled The Development of an Instrument to Measure Mental Toughness in Sport and the work presented in the thesis are both my own, and have been generated by me as the result of my own original research. I confirm that:

- this work was done wholly or mainly while in candidature for a research degree at this University;
- where any part of this thesis has previously been submitted for a degree or any other qualification at this University or any other institution, this has been clearly stated;
- where I have consulted the published work of others, this is always clearly attributed;
- where I have quoted from the work of others, the source is always given. With the exception of such quotations, this thesis is entirely my own work;
- I have acknowledged all main sources of help;
- where the thesis is based on work done by myself jointly with others, I have made clear exactly what was done by others and what I have contributed myself;
- parts of this work have been published as: [please list references]

Signed:

Date: 13/01/2014

LIST OF PUBLICATIONS FROM THESIS

Conference Contributions, refereed;

Birch, P. D. J., Greenlees. I. A., Lowry, R. G., & Coffee, P. (2012). The Mental Toughness Questionnaire-48: A factorial analysis. Paper presented at the British Psychological Society Annual Conference, UK. April, 2012.

ACKNOWLEDGEMENTS

First and foremost I would like to thank my Director of Studies, Dr Iain Greenlees for his support, patience and direction over the past four years and my time as an undergraduate and postgraduate student at the University of Chichester. His enthusiasm, personal integrity, and most importantly, his faith in my abilities were invaluable to maintain my focus and motivation throughout my studies and PhD tenure. I would like to thank my co-supervisor Dr Ruth Lowry for her support and advice in bridging the gap between psychology theory and statistics. Her patience and expertise within the area of Structural Equation Modelling was an invaluable resource which undoubtedly facilitated my understanding of the area. I would also like to thank my external supervisor Dr Peter Coffee who provided me with impartial advice and constructive criticism which undoubtedly enhanced the quality of my research. Both consciously and subconsciously, they have taught me how rigorous science is done.

I am grateful to my PhD colleagues and friends; Helen Barrett, John Batten, Murray Cod, Sean Figgins, Carla Gallagher, James Gavin, Cheng-Wei Li, Jordan Matthews, Pete Mundy, Russell Peters, Laura Shafe, Matt Sitch, Andy West, and David Williams who provided a community and made my experience of studying at the University of Chichester all the more enjoyable. Not forgetting my good friend Paul Taylor who always provided me with a social outlet when things were tough. I would also like to thank Dr Philip Kearney, Dr Jenny Page, and Dr Matthew Smith for sharing their PhD experiences with me. I am grateful to Simon Crampton who provided me with a critical yet friendly perspective on all things related to mental toughness and statistics. Sharing similar frustrations and successes was an invaluable resource to my development as a researcher. I also extend my gratitude to Dr Daniel Gucciardi who helped me keep up-to-date with research developments in mental toughness.

I owe gratitude to Dr Beverly Hale whose door was always open and was ever willing to impart guidance and clarification regarding all things statistical. Many thanks to the administrative staff of the Research and Employer Engagement Office (REEO) for allowing me to attend and present at international and national conferences; this was instrumental in my progression from MPhil to PhD. This thesis would not have been possible were it not for those dedicated participants who invested their time into this project. Although there are too many to mention, their time and effort will not be forgotten.

I would like to thank my family, especially my parents Dolores and Dave Birch, my sister Angela, my Nan Ruth Birch, and my Aunty and Uncle Anne and Graham Curtis. Without their moral, financial and emotional support I would have never made it this far. I would also like to thank my close friends, especially Luke Stewart, Jack Wade, and John Wilson for providing me with the support and social outlet so that I could take a break from the trials and tribulations of the research process. Furthermore I would like to thank Sara, Paul, Jack and Natalie Nicholson for their support throughout the PhD. Last but certainly not least, I would like to thank my loving girlfriend Emma Nicholson. Her understanding of sports science coupled with her caring and encouraging attitude provided me with the everyday support which was invaluable to my progress throughout the PhD.

CHAPTER 1.0

INTRODUCTION

Since 2002, researchers have developed a number of models to examine Mental Toughness (MT) in sport. To date, Clough, Earle, and Sewell's (2002) 4C's model of MT and Earle's (2006) 6C's model of MT have received the most research attention (Gucciardi, Hanton, & Mallett, 2012). Clough and colleagues' 4/6C's model of MT proposes that MT resembles tenets outlined in hardiness theory (Kobasa, 1979) where a combination of traits (challenge, commitment, control) are thought to motivate one to appraise potentially threatening situations as opportunities for personal development (e.g., Kobassa, 1979; Maddi, 2002; Maddi & Kobassa, 1984). However, Clough and colleagues suggest that hardiness alone does not fully encapsulate MT and added a fourth trait, confidence, to account for the physical and mental demands of competitive sport. According to Clough and colleagues, confidence is an important factor relating to sport which is not considered as a distinct facet of hardiness theory. Clough *et al.'s* 4C's model of MT therefore conceptualises MT as a combination of four traits; challenge, commitment, control, and confidence. Despite advocating the 4C's model of MT, doctoral research by Earle postulated that MT is best understood by a 6C's model whereby control is subdivided into two nested components relating to control – emotion and control life and whereby confidence is subdivided into two nested components relating to confidence – abilities and confidence – interpersonal. Given the extensive body of literature within mainstream psychology which supports the stress buffering (e.g., Bartone, 1999; Sheard & Golby, 2006) and performance enhancing (e.g., Judkins & Rind, 2005; Maddi *et al.*, 2006) role of the traits thought to make-up hardiness, it has been proposed (Clough *et al*; Earle) that the 4/6C's models of MT may provide a conceptually well-grounded approach to examine the traits thought to underpin MT.

In order to fully test the applicability and validity of the 4/6C's model of MT, Earle (2006) developed the Mental Toughness Questionnaire-48 (MTQ48). The MTQ48 is a 48-item inventory assessing dispositional MT. In the MTQ48's development, Earle offered support for its convergent validity by reporting correlations with a number of related psychological constructs including optimism, self-image, life satisfaction, self-efficacy, and stability. In addition, Clough *et al.* (2002) offered support for the criterion validity of the MTQ48 by reporting two studies whereby MTQ48 scores differentiated participants on perceived exertion in a cycling task and performance on a cognitive task following negative feedback. Research using the MTQ48 has predominantly examined the cognitive correlates of MT by correlating scores from the MTQ48 with a range of coping related variables such as coping style (e.g., Nicholls, Polman, Levy, & Blackhouse, 2008), coping effectiveness (e.g., Kaiseler, Polman, & Nicholls, 2009), and coping self-efficacy (e.g., Nicholls, Levy, Polman, & Crust, 2011). Research has also correlated scores from the MTQ48 with other related psychological variables such as leadership preference (Crust & Azadi, 2009), psychological skill usage (Crust & Azadi, 2010), risk taking (Crust & Keegan, 2010), and dispositional flow (Crust & Swann, 2011a).

Despite this evidence base, many researchers (e.g., Connaughton & Hanton, 2009; Gucciardi *et al.,* 2011) have questioned the rigour underpinning the MTQ48's development. Close inspection of the procedures used to develop the MTQ48 reveals a number of concerns relative to the recommendations outlined in the scale development literature (e.g., Clark & Watson, 1995; Hinkin, 1995; MacKenzie, Podsakoff, & Podsakoff, 2011). First, there appears to be concerns with the definitions forwarded by Clough *et al.* (2007) to underpin the 4/6C's model of MT in that only characteristics are provided as opposed to clear and concise definitions (Churchill, 1979; Hinkin; MacKenzie *et al.*). Second, inspection of MTQ48 item content reveals that a number of items may be poor representations of their hypothesised factors which may limit their ability to measure what they intend capture (MacKenzie *et al.*). Third, Earle did not use an established method to assess item content validity such as the Content Validation Index (CVI; Lynn, 1986). Finally, inspection of the MTQ48's validation protocols reveals that there was a reliance on Principle Components Analysis (PCA) to examine its factorial validity. According to the scale development literature (e.g., MacKenzie *et al*; Marsh, 2007), statistical techniques such as Confirmatory Factor Analysis (CFA) are required to test the robustness of an *a priori* model structure. Consequently, it could be argued that the scale development practices adopted by Earle (2006) reduces the validity of the MTQ48 and subsequently explains why it has received criticism.

The lack of research using CFA to support the MTQ48's factorial validity has led some researchers (e.g., Connaughton & Hanton, 2009; Gucciardi *et al.*, 2011) to question its adequacy in measuring the 4/6C's model of MT. Researchers using the MTQ48 (e.g., Crust & Swann, 2011a, 2011b) have argued that its factorial validity was independently supported by Horsburgh, Schermer, Veselka, and Vernon (2009). However, close inspection of Horsburgh *et al.'s* study reveals a number of limitations. The most pronounced limitation is the absence of any empirical data (i.e., fit indices, parameter estimates) to support their conclusion about the superiority of the first-order four factor solution when compared with a unidimensional single-factor solution (Gucciardi *et al.*, 2012). The absence of the relevant Exploratory Factor Analysis (EFA) and CFA statistics therefore makes it difficult for the reader to draw any robust conclusions regarding the adequacy of the first-order four factor model of the MTQ48. Research by Gucciardi and colleagues has, to date, provided the only study to rigorously test the factorial validity of the first-order four factor model of the MTQ48 in two independent samples; a) athletes, and b) full-time employees. The preliminary evidence yielded by Gucciardi *et al.* provided little support for the factorial validity of the MTQ48 suggesting that there could be problems regarding its adequacy to measure the 4/6C's model of MT. However, there are a number of limitations with Gucciardi *et al.'s* (2012) study. For instance, Gucciardi *et al.* did not report a factorial analysis of the hypothesised six factor model and did not report second-order representations of the four and six factor models of the MTQ48. Consequently, there is currently no information in the literature as to the adequacy of the six factor model structure and the adequacy of second-order representations of the four and six factor models of the MTQ48. Further research is therefore warranted to provide a more comprehensive examination of the MTQ48's factorial validity.

Given the concerns regarding the development of the MTQ48 and the evidence relating to its factorial validity, there appears to be a degree of ambiguity surrounding the utility of the measure. Although the preliminary findings of Gucciardi *et al.* (2012) provide little support for the utility of the MTQ48, scale development and validation is an ongoing process which requires replication and verification of findings (e.g., Marsh, 1997, 2002, 2007). Using the scale development literature (e.g., Clark & Watson, 1995; MacKenzie *et al.,* 2011) as a guiding framework, further research is therefore warranted to extend Gucciardi *et al.'s* research by examining the factorial validity of the first- and second-order models of the four and six factor models of the MTQ48 and, where appropriate, to develop a robust measure of MT as defined by Clough and colleagues. The provision of a valid tool to measure the 4/6C's model of MT will not only enable researchers to examine the traits thought to underpin MT, but it may also provide a psychometrically sound tool to measure MT. This is particularly important given that Gucciardi *et al.'s* (2011) recent review of existing MT measures concluded that "in assessing the adequacy of each mental toughness inventory on conceptual, statistical or empirical, and practical grounds, we conclude that, at present, no measure sufficiently satisfies all three criteria" (p. 128). The aim of this thesis, therefore, was to provide a comprehensive examination of the validity and reliability of the MTQ48 in an effort to provide a valid measure of MT.

1.1 An overview of the thesis

Having reviewed the theories and research concerning the MTQ48 in Chapter 2, Chapter 3 provides a factorial analysis of the MTQ48 using a large sample of competitive student athletes. As this study found little support for the factorial validity of the MTQ48, further research was required to develop and examine the factorial validity of an adapted MTQ48 in an effort to provide a valid instrument to measure the 4/6C's model of MT. Chapter 4 presents a study detailing the development and factorial analysis of an adapted questionnaire, the University of Chichester Mental Toughness Questionnaire (UCMTQ), which incorporates regenerated items to better represent the 4/6C's model of MT. Given the lack of support found for the factorial validity of the UCMTQ, Chapter 5 provides a study detailing the development and factorial analysis of a further adapted questionnaire, the Hardiness Confidence Questionnaire (HCQ), which incorporates redefined components of the 4C's model of MT and regenerated items to better represent the traits underpinning the model. Chapter 6 then provides a study detailing the construct validity of the Revised Hardiness Confidence Questionnaire (HCQ-R) whereby the convergent and predictive validity and test-retest reliability of the instrument are examined. Chapter 7 provides a general discussion of the results reported in Chapters 3 to 6, highlighting the collective findings and implications of the four studies, as well as providing suggestions for future avenues of research investigating the measurement of MT in sport.

CHAPTER 2.0

REVIEW OF LITERATURE

2.1 Mental Toughness in Sport

MT is a term that has been used since the 1950's by athletes, coaches, and the media to describe superior mental characteristics of successful athletes (Gucciardi, Gordon, & Dimmock, 2008). Researchers have attempted to explain MT in a variety of manners that include MT as a defence mechanism against adversity (Alderman, 1974; Favret & Benzel, 1997; Gould, Hodge, Peterson, & Petlichkoff, 1987; Loehr, 1982, 1986; Tapp, 1991), a critical asset for athletes to endure the long hours of strenuous training associated with top-level performance (Bull, Albinson, & Shambrook, 1996; Goldberg, 1998), and a decisive factor in distinguishing successful and unsuccessful performances (Loehr, 1982, 1986, 1995; Luszki, 1982). Although these initial discussions provided the first insight into the concept of MT, the conceptual underpinnings of the proposed definitions of MT were questionable (i.e., the vast literature was based on coaching/counselling experience and anecdotal evidence as opposed to using established qualitative techniques), and many of the studies investigated mental skills rather than MT *per se* (Connaughton, Thelwell, & Hanton, 2012). The lack of empirical research exacerbated misinterpretation and created confusion regarding a clear understanding of MT (Connaughton & Hanton, 2009). In 2001, Fourie and Potgeiter were the first researchers to empirically examine MT in sport. Interviews with expert coaches and elite athletes from a variety of sports yielded 12 components of MT: Motivation level, coping skills, confidence maintenance, cognitive skill, discipline and goal directedness, competitiveness, possession of prerequisite physical and mental requirements, team unity, preparation skills, psychological hardiness, religious convictions, and ethics. Fourie and Potgeiter's (2001) study stimulated interest into the construct of MT and was the catalyst for an array of research programs designed to overcome the anecdotal issues inherent in the early MT literature.

To date, two schools of thought have emerged within the literature. The first school of thought, led by Jones, Hanton, and Connaughton (2002) and Gucciardi *et al.* (2008) has primarily used qualitative methods to define MT and its key components in various sports from the perspectives of athletes, coaches, parents, and sport psychologists. In an effort to provide conceptual clarity and to overcome the atheroretical approaches used in previous work examining MT, this school has utilised Kelly's (1991) Personal Construct Psychology (PCP) coupled with qualitative methods (e.g., semi-structured interviews) to underpin the examination of MT. Jones *et al.* and Gucciardi *et al.* argue that PCP offers a framework whereby an individual's views, experiences, meanings, and perceptions can be understood clearly and accurately which allows for a more comprehensive examination of the MT phenomenon. To date, six studies have examined the key components of MT (Bull, Shambrook, James, & Brooks, 2005; Coulter, Mallett, & Gucciardi, 2010; Gucciardi *et al*; Jones *et al.*, 2002; Jones, Hanton, & Connaughton, 2007; Thelwell, Weston, & Greenlees, 2005). These six studies have yielded an overwhelmingly large number of characteristics thought to make-up the mentally tough performer. However, whilst 118 attributes have been identified, Connaughton *et al.* (2012) suggest that the identified attributes can be broadly classified under nine major themes: Belief, coping/handling pressure, focus/commitment, motivation, control, sporting intelligence/knowledge, tough/resilient attitude, personal values, and physical toughness.

Jones *et al.* (2002, 2007) interviewed international level athletes, coaches and sport psychologists to establish a definition of MT and to identify the essential attributes of the mentally tough performer. Jones *et al.* (2002) defined MT as:

> "Having the natural or developed psychological edge that enables you to: Generally, cope better than your opponents with the many demands (competition, training, lifestyle) that sport places on a performer. Specifically, be more consistent and better than your opponents in remaining determined, focused, confident, and in control under pressure" (p. 209).

In a similar vein, research by Gucciardi *et al.* (2008) examined the context specific nature of MT by interviewing elite level Australian football coaches to determine a definition of MT and the key characteristics which make-up the mentally tough Australian football player. Gucciardi *et al.* defined MT as: "A collection of values, attitudes, behaviours, and emotions that enable you to persevere and overcome any obstacle, adversity, or pressure experienced, but also to maintain concentration and motivation when things are going well to consistently achieve your goals" (p. 218). Although the respective MT definitions in both sport general and sport specific contexts provided the first operational definitions of MT (see also Bull *et al.,* 2005; Thelwell *et al.,* 2005), one may question the clarity of the resultant definitions relative to whether MT is thought to be a trait or a developed construct. Inspection of Jones *et al.'s* (2002) definition reveals that MT is thought to be a natural (trait) or developed construct whereas Gucciardi *et al.'s* definition does not clearly state whether MT is thought to be a trait or a learned construct. The apparent ambiguity surrounding these definitions makes it difficult to specify what MT actually constitutes at a conceptual level. Researchers and practitioners may be confused regarding how best to conceptually embed their ideas of MT given that there is no clear distinction of whether MT is underpinned by learning and/or by personality. However, the second school of thought appears to provide clearer theoretical assertions regarding the make-up of MT.

2.2 Dispositional School of Thought of Mental Toughness

The second school of thought, led by Clough *et al.* (2002) and Earle (2006) have proposed that MT is best understood as a personality trait whereby a combination of dispositions are thought to underpin MT. Clough and colleagues argued that theoretical models within the health psychology literature offer the practitioner an insight into the components of MT, but are limited in their application to identify the unique nature of the physical and mental demands of competitive sport. Clough and colleagues further argued that ‘common sense’, yet ecologically valid perspectives of sport psychology practitioners may enable researchers to better identify and understand sport-specific elements of MT, but lack the theoretical rigour needed to guide empirical research. In light of these arguments, Clough and colleagues suggested that MT resembled tenets outlined in hardiness theory (Kobasa, 1979) where a combination of dispositions (challenge, commitment, control) are thought to motivate one to appraise potentially threatening situations as opportunities for personal development (e.g., Kobassa, 1979; Maddi, 2002; Maddi & Kobassa, 1984). Clough and colleagues used information reported in the hardiness literature to define each subcomponent of Kobasa’s (1979) hardiness theory. Specifically, Clough *et al.* (2002, p. 35) and Earle (2006, p.60) stated that challenge is expressed as “the belief that change, rather than stability, is normal in life and that the anticipation of changes are interesting incentives to growth rather than threats to security” (Kobasa, Maddi, & Khan, 1982). Commitment is a “tendency to involve oneself in, rather than experience alienation from whatever one is doing, or encounters” (Maddi, Hoover, & Kobasa, 1982). Control is expressed as a “tendency to feel and act as if one is influential (rather than helpless) in the face of the varied contingencies of life” (Averill, 1973; Kobasa *et al.*, 1982; Seligman, 1975).

According to Kobassa (1982), challenge leads individuals to anticipate and welcome new experiences and to hone their responses to the unexpected so that situations are perceived as stimulating, rather than as threatening. Commitment reduces the perceived threat of stressful situations by providing individuals with a sense of purpose and the belief that they can turn to others and others can turn to them to deal with stress (Kobassa). Control provides an individual with a large variety of responses with which to deal with stressful situations, which in turn, enables them to choose the best way of coping (Kobassa). Maddi and Kobasa (1984) argued that challenge, commitment, and control diminishes stress and strain, and affords the enhancement of performance and health. Despite there being an impressive body of literature to support the stress buffering (e.g., Bartone, 1999; Sheard & Golby, 2006) and performance enhancing (e.g., Judkins & Rind, 2005; Maddi *et al.*, 2006) role of hardiness in non-sporting populations, relatively few studies have examined its impact within sport. Research examining hardiness in sport has found it to be positively related to a variety of variables including adaptive coping strategies (Goss, 1994; Wadey, Evans, Hanton, & Neil, 2011), injury time-loss (Ford, Eklund, & Gordon, 2000), standards of competition (Golby & Sheard, 2004; Sheard, 2009) and performance (Maddi & Hess, 1992). In addition, research has found hardiness to be negatively related to negative mood states (Goss; Prapavessis & Grove, 1994; Skirka, 2000), undesirable post-injury psychological responses (Wadey *et al.*), maladaptive coping (Goss; Wadey *et al.*), and perceived stress and burnout (Kelley, 1994; Kelley, Eklund, & Ritter-Taylor, 1999; Martin, Kelley, & Dias, 1999; Martin, Kelley, & Eklund, 1999; Skirka, 2000).

However, Clough *et al.* (2002) and Earle (2006) proposed that hardiness does not fully encompass MT in that it does not account for the physical and mental demands of competitive sport. Clough and colleagues argued that confidence is an important factor relating to sport performance which is not considered in existing hardiness models. Despite Clough and colleagues positing that MT is made up of a combination of four dispositions (challenge, commitment, control, confidence), two marginally different models were developed. Specifically, Clough *et al.* argued that MT is a combination of hardiness and confidence which resulted in the conception of the 4C's model of MT (challenge, commitment, control, confidence). Clough *et al.* stated that evidence to support the ecological validity of the 4C's model of MT was obtained from the views of a Rugby league team whereby confidence was identified to be an important construct relating to performance. Although Clough *et al.* clearly defined challenge, commitment, and control in line with the hardiness literature, no definitions were provided to underpin the 4C's model of MT. Nonetheless, Clough *et al.* did provide some information to describe the mentally tough performer:

"Mentally tough individuals tend to be more sociable and outgoing; as they are able to remain calm and relaxed, they are competitive in many situations and have lower anxiety levels than others. With a high sense of self-belief and an unshakable faith that they can control their own destiny, these individuals can remain unaffected by competition or adversity" (p. 38).

Despite supporting the contention that MT is a combination of challenge, commitment, control, and confidence, Earle (2006) posited that MT is best understood when the control and confidence constructs are subdivided into two nested components. Specifically, Earle conducted 12 interviews with a variety of sports people to explore the make-up of MT. These included three rugby coaches, one rugby chief executive, two rugby players, two golfers, two footballers, and two squash players. Although findings showed that most themes could be categorised under the construct of confidence and Kobasa's (1979) model of hardiness, Earle suggested that control and confidence have a more complex structure. Earle stated that control (control - emotion, control - life) and confidence (confidence - ability, confidence - interpersonal) should incorporate two nested components. This resulted in the 6C's model of MT. Although Earle did not provide definitions to underpin the 6C's model of MT, later work by Clough, Marchant, and Earle (2007) provided definitions to underpin the respective constructs of the 4/6C's model (see Table 2.1).

Table 2.1. Definitions underpinning the 4/6C's model of MT.

Factor	Construct definitions
Challenge (sometimes referred to as change orientation)	"Describes the extent to which individuals see challenges as opportunities. Individuals who see them as opportunities will actively seek them out and will identify problems as ways for self-development. At the other end challenges are perceived as problems and threats. So, for example, at one end of the scale we find those who prefer to minimise their exposure to change and the problems that come with that – and will strongly prefer to work in stable environments".
Commitment (sometimes referred to as 'stickability')	"Sometimes described as 'stickability', this describes the ability for an individual to carry out tasks successfully despite any problems or obstacles that arise whilst achieving the goal. Consequently an individual who scores at the high end of the scale will be able to handle and achieve things to tough unyielding deadlines. Whereas an individual at the other end will need to be free from those kind of demands to achieve their goals".
Control	"Individuals who score high on this scale feel that they are in control of their work and of the environment in which they work. They are capable of exerting more influence on their working environment and are more confident about working in complex or multi-tasked situations. This means for example that, at one end of the scale individuals are able to handle lots of things at the same time. At the other end they may only be comfortable handling one thing at a time. Ongoing development of MTQ48 has enabled the identification of 2 subscales to this scale:"
Control - emotion	"Individuals scoring highly on this scale are better able to control their emotions. They are able to keep anxieties in check and are less likely to reveal their emotional state to other people".
Control - life	"Individuals scoring higher on this scale are more likely to believe that they control their lives. They feel that their plans will not be thwarted and that they can make a difference".
Confidence	"Individuals who are high in confidence have the self-belief to successfully complete tasks, which may be considered too difficult by individuals with similar abilities but with lower confidence. Less confident individuals are also likely to be less persistent and may make more errors. For example, individuals at one end of the scale will be able to take setbacks (externally and self-generated) in their stride. They keep their heads when things go wrong and it may even strengthen their resolve to do something. At the other end individuals will be unsettled by setbacks and will feel undermined by these. Their heads are said to 'drop'".
Confidence - abilities	"Individuals scoring highly on this scale are more likely to believe that they are a truly worthwhile person. They are less dependent on external validation and tend to be more optimistic about life in general".
Confidence - interpersonal	"Individuals scoring highly on this scale tend to be more assertive. They are less likely to be intimidated in social settings and are more likely to push themselves forward in groups. They are better able to cope with difficult or awkward people".

Adapted from Clough *et al.* (2007).

Given the extensive body of evidence to support the stress-buffering (e.g., Bartone, 1999; Sheard & Golby, 2006) and performance enhancement effects (e.g., Judkins & Rind, 2005; Maddi *et al.*, 2006) of hardiness in non-sporting contexts and the preliminary evidence to support its impact in sporting contexts (e.g., Hanton *et al.*, 2003; Maddi & Hess, 1992), further research is warranted to investigate the 4/6C's model of MT to determine the impact of these traits within the context of sport. Research examining the traits thought to underpin MT could enhance our understanding of their impact upon sporting performance. The role of personality in predicting athletic success has received ample research attention (e.g., Aidman, 2007; Egloff & Gruhn, 1996; Evans & Quarterman, 1983; Gee, Marshall, & King, 2010; Morgan & Johnson, 1978; Morgan, O'Connor, Ellickson, & Bradley, 1988; Piedmont, Hill, & Blanco, 1999). Despite researchers (e.g., Aidman; Vealey, 1989) criticising the use of descriptive and cross-sectional designs to examine the short-term impact of traits upon sporting performance, the predictive contributions of personality traits when examined in a longitudinal manner are thought to be more valid and reliable (e.g., Hogan, 1998; Hogan & Shelton, 1998; Morris, 1995; Pervin, 1996). Research using longitudinal designs (e.g., Aidman; Gee *et al.*) has found personality traits (e.g., competitiveness, openness) to be an effective means of predicting who ultimately succeeds at the senior level of competition when assessed seven and fifteen years later, respectively. Gee and colleagues have argued that the assessment of personality appears to add to a practitioner's (e.g., coaches, sport psychology consultants) ability to predict an athlete's longitudinal athletic potential. If the traits that underpin MT can be established, research can examine the predictive contribution of these MT traits and assess their efficacy in longitudinally assessing sporting success.

2.3 Research Examining the 4/6C's Model of Mental Toughness in Sport

The 4/6C's model of MT has received the most research attention within sport and non-sport contexts (Gucciardi *et al.*, 2012). Although research has been conducted in non-sporting contexts including the military (Simpson, Gray, & Florida-James, 2006), police recruits and students (Clough *et al.*, 2007), and business (Marchant *et al.*, 2008), the focus of this review of literature will be the application to sporting contexts.

2.3.1 Comparative Analyses of Mental Toughness

Research has investigated differences in MT across a range of demographic variables in sport. Nicholls, Polman, Levy, and Blackhouse (2009) found that male athletes scored significantly higher than female athletes on challenge, control – emotion, control – life, and confidence abilities. Similarly, Crust and Keegan (2010) revealed that male student athletes reported significantly higher levels of total MT and confidence – abilities than female athletes. However, research by Crust (2009) and Crust and Azadi (2010) did not find statistical differences across gender. Although research by Crust and Azadi found that county level athletes reported significantly higher levels of MT than club or university athletes, other research has failed to show differences in MT across different achievement levels, sport type, age, and playing experience (e.g., Crust, 2009; Nicholls *et al.*, 2009). Consequently, the respective comparative analyses appear to be equivocal.

2.3.2 Cognitive Correlates of Mental Toughness

In light of the 4/6C's model conceptualising MT as a combination of dispositions, many researchers have examined the relationship between MT and a number of related traits. Crust and Keegan (2010) identified that qualitative evidence suggests that risk taking could be an important facet of MT (Bull *et al.*, 2005) and research examining risk taking (e.g., Llewellyn & Sanchez, 2008) could suggest that specific elements of MT (i.e., challenge and confidence in one's ability) might predict an individual's tendency to risk-taking. Crust and Keegan found that although physical risk-taking attitudes were positively and significantly related to total MT, commitment, and confidence – abilities, only confidence - interpersonal was positively and significantly related to psychological risk-taking attitudes. Although Crust and Keegan provided some rationale to underpin the hypothesised link between MT and risk-taking, very little explanation was offered as to whether risk taking is a facet of MT or a behaviour demonstrated by those who are mentally tough.

Crust and Swann (2011a) found MT to be positively and significantly related to dispositional global flow. Crust and Swann stated that these findings are consistent with previous research which has found confidence to be positively and significantly related to global flow (Hodge, Lonsdale, & Jackson, 2009) and propositions arguing that a balance between perceived high ability and challenging situations will facilitate flow (Jackson & Kimiecik, 2008). Research by Kaiseler *et al.* (2009) and Nicholls *et al.* (2008, 2011) found MT to be positively and significantly related to optimism and negatively and significantly related to pessimism. Despite the emerging relationships with other traits, little is known about the nature of these relationships.

Research has also examined the relationships between MT and a variety of other psychological variables. Crust and Azadi (2009) examined the relationship between MT and leadership preference and found MT to be positively and significantly related to training and instruction preference. Crust and Azadi (2010) also correlated scores yielded from the MTQ48 and the Test of Performance Strategies (TOPS; Thomas, Murphy, & Hardy, 1999) and found MT to be positively associated with the self-talk, emotional control, and relaxation subscales. Further results showed that the commitment subscale exhibited the greatest number of statistically significant correlations with psychological skill usage (13 of the 16 subscales of the TOPS inventory). Crust and Azadi argued that the dominance of commitment in relation to psychological skill usage may be a result of one's motivation to be deeply involved in sport and thus lead an individual to seek out alternative ways of enhancing performance.

Noting that the early qualitative research (e.g., Fourie & Potgieter, 2001; Jones *et al.* 2002) emphasised the importance of the mentally tough performer being able to cope with pressure, a series of research studies have examined the relationship between MT and a number of coping related variables. In general, findings show MT to be positively and significantly related to the use of problem/approach coping strategies and coping effectiveness, and negatively related to emotional and avoidance coping strategies (Nicholls *et al.*, 2008, 2011; Kaiseler *et al.*, 2009). The aforementioned studies illustrate the relative abundance of research which has examined the cognitive correlates of the 4/6C's model of MT. However, the cross-sectional designs used in these studies may limit the predictive power of the findings.

2.3.3 Affective Correlates of Mental Toughness

In an effort to extend our understanding of the relationship between MT and coping, Crust (2009) investigated the affective consequences of MT. Findings showed no significant relationships between affect intensity and MT. Crust concluded that these findings suggest that there is no relationship between MT and the intensity of emotions which one experiences. Consequently, it appears that differences in coping ability (e.g., Nicholls & Polman, 2007) and/or the interpretation of anxiety (e.g., Jones, Hanton, & Swain, 1994; Jones & Swain, 1995) are more accurate markers of an individual's ability to cope with stress as opposed to the intensity of emotions one experiences.

2.3.4 Behavioural Correlates of Mental Toughness

Despite there being a need to examine a constructs affective and cognitive correlates, some researchers (e.g., Andersen, McCullagh, & Wilson, 2007) have placed greater importance on establishing the link between psychological constructs and actual behaviour. Specifically, Andersen *et al.* argue that unless self-report scores are related back to overt behavioural variables (e.g., performance), one cannot truly understand the meaning of the scores obtained. Clough *et al.* (2002) offered two studies to support the behavioural correlates of MT. The first study examined the relationship between MT and perceived effort. Participants completed the MTQ48 and were classified (via a median split) as either possessing high or low levels of MT. Participants were required to cycle at various workload levels (30%, 50%, or 70%) relative to their $VO_{2\,max}$ for a period of 30 minutes. Findings revealed that although there were no differences in perceived exertion at the 30% workload level, participants who were high in MT reported significantly lower levels of perceived exertion at the 70% workload level than participants low in MT.

In the second study, Clough *et al.* (2002) investigated the effect of feedback on performance relative to low and high levels of MT. Participants were either given positive or negative feedback following the completion of a number of motor tasks. In order to assess participants' responses to the respective feedback received, participants were then required to complete a cognitive planning task. Findings revealed a significant interaction ($F = 4.36$, $p < 0.05$) between MT and feedback in that participants high in MT performed the cognitive planning task consistently irrespective of the feedback received, whereas participants low in MT performed significantly worse following negative feedback. Although the evidence reported by Clough *et al.* appears to provide preliminary support for the behavioural correlates of the 4/6C's model of MT, the absence of information regarding the procedures administered (e.g., measures used, tasks used, data analyses conducted), participants used (e.g., demographics of participants), and the comprehensive presentation of statistics makes it difficult for the reader to draw any robust conclusions from the findings presented. Furthermore, inspection of Earle's (2006) unpublished doctoral research shows that when these studies were replicated, findings generally showed no significant main effects and/or interaction effects between MT and the respective outcome measures tested.

Noting the need to further examine the behavioural correlates of MT, Crust and Clough (2005) examined the relationship between MT and performance by examining the relationship between MT and physical endurance. Participants were asked to lift a dumbbell weighing approximately 1.5% of their body weight at a 90 degree angle between their arm and torso for as long as possible. Findings revealed significant and positive relationships between endurance time and total MT ($r = 0.34$, $p < 0.05$), control ($r = 0.37$, $p < 0.05$) and confidence ($r = 0.29$, $p < 0.05$), but not for challenge ($r = 0.22$) or commitment ($r = 0.23$). Given the limitations associated with the research which has examined the behavioural correlates of MT, one may question the extent to which MT scores are related back to actual behaviours. Consequently, it could be argued that the data gleaned from these studies are limited in providing support for the behavioural correlates of MT (Andersen *et al.*, 2007).

Upon evaluating the research which has examined the 4/6C's model of MT, it is clear that researchers have placed greater importance on examining the cognitive correlates of MT as opposed to examining the affective and behavioural correlates. Consequently, it could be argued that the unbalanced research attention given to the respective correlates of MT has limited our understanding of the 4/6C's model of MT. Indeed, Andersen *et al.* (2007) suggest that adopting a triangulation approach to encompass all correlates may facilitate our understanding of the individuals we study and the construct under examination.

2.4 Criticisms of Clough and Earle's Approach to Mental Toughness

2.4.1 Conceptual Issues

Although a growing number of researchers advocate the use of the 4/6C's model of MT as a valid conceptualisation of MT in sport, researchers (e.g., Connaughton & Hanton, 2009; Gucciardi *et al.*, 2011) have questioned Clough *et al.'s* (2002) and Earle's (2006) rationale to conceptualise MT as a combination of hardiness and confidence given that no detailed information was provided to underpin the associations of MT to hardiness theory. Moreover, researchers (e.g., Connaughton & Hanton; Gucciardi *et al.*) have raised concerns regarding the scientific rigour underpinning the 4C's model of MT given that no information was presented regarding data collection and analysis of the views of the key stakeholders (i.e., sport psychology practitioners, Rugby league team) used to develop the model. However, Earle's doctoral research has been overlooked in the literature which appears to provide information regarding the development of the 6C's model of MT. Earle used information gleaned from interviews with 12 sports people to inform the development of a model to conceptualise MT. Despite Clough *et al.'s* failure to provide detailed information regarding the development of the model and the associations between hardiness and confidence, the work of Earle appears to provide qualitative information to suggest that the components included in the 4/6C's model of MT are conceptually justified. Indeed, the well-established importance of self-belief in the make-up of MT (e.g., Jones *et al.*, 2002; Gucciardi *et al.*, 2008) and the well-established relationship between confidence and sporting performance (e.g., Jones & Hanton, 2001; Vealey, 2001) provides evidence to support the integration of confidence in a model to conceptualise MT.

Furthermore, the extensive body of literature which supports the stress buffering (e.g., Bartone, 1999; Sheard & Golby, 2006) and performance enhancing (e.g., Judkins & Rind, 2005; Maddi *et al.*, 2006) effects of hardiness in non-sporting contexts and the preliminary evidence to support its impact in sporting contexts (e.g., Hanton *et al.*, 2003; Maddi & Hess, 1992) provides supporting evidence to embed MT in hardiness theory. Therefore, the tenets outlined in the 4/6C's model of MT appear to provide a conceptually well-grounded approach to examine the traits thought to underpin MT. Indeed, the application of hardiness theory has been noted as a key strength of Clough *et al.'s* 4C's model of MT (Gucciardi *et al.*, 2011). Given the potential utility of the 4/6C's model of MT, one may suggest that researchers may be prematurely dismissing the 4/6C's model of MT as a valid model to examine MT. Further research is therefore warranted to examine the 4/6C's model of MT.

2.4.2 Measurement issues

Despite the potential utility of the 4/6C's model of MT, there appears to be number of problems relating the MTQ48. Many researchers (e.g., Connaughton & Hanton, 2009; Gucciardi *et al.*, 2011) have questioned the adequacy of the MTQ48 to measure the 4/6C's model of MT given that in its conception, Clough *et al.* (2002) offered no evidence for its factorial validity. However, very little attention has been given to Earle's (2006) research which provides more comprehensive information regarding the development and validation of the MTQ48. According to Earle, PCA with varimax rotation was used to examine the adequacy of an initial pool of 66 items designed to measure the 4/6C's model of MT. A total of 215 participants (male = 123, female = 92) from a variety of occupations including students (*n* =129), professional athletes (*n* = 52), and administration/managerial staff were asked to complete the 66-item measure. The age of the sample ranged from 18 to 57 years (*M* = 24.65, *S.D.* = 8.35). Eigenvalues greater than one were accepted. Inspection of the initial solution revealed that 18 factors had eigenvalues greater than one accounting for 64.5% of the accumulative variance. Despite this finding, Earle inspected four solutions ranging from four to seven factors. On the basis that the six factor solution (38.9%) accounted for a greater percentage of accumulative variance than the four (28%) and the five (30.1%) factor solutions, Earle extracted a six factor model of MT which incorporated 48-items. This resulted in the emergence of the MTQ48. Table 2.2 provides an analysis summary of the extracted six factor solution.

In an effort to provide support for the extracted six factor model of the MTQ48, Earle (2006) re-examined the factor structure of the MTQ48. Despite Earle providing limited information for these secondary analyses, Clough *et al.* (2007) provided more comprehensive information in the MTQ48 Technical Manual. A total of 963 people (men = 338, women = 376, missing = 249) from a variety of occupations including students (*n* = 619), administrators/managers (*n* = 136), engineers (*n* = 42), and athletes (*n* = 166) were asked to complete the MTQ48. The age of the sample ranged from 18 to 59 years (*M* = 24.21, *S.D.* = 5.23). Similarly, PCA with a varimax rotation was used and eigenvalues greater than one were accepted. Table 2.3 provides an analysis summary of the extracted six factor solution.

Despite Clough *et al.* (2007) and Earle (2006) concluding that the respective analyses provided support for the six factor model of the MTQ48, one may question these conclusions. Close inspection of the factor loadings of the respective PCA's suggest that the MTQ48 items may not be accurate representations of their respective factors given that the lowest values could be considered too low for interpretation (< 0.32; Comrey & Lee, 1992) and the highest values could not be classified as excellent (> 0.71; Comrey & Lee). Clough and colleagues only report the ranges of factor loadings for each respective factor which prevents the reader from ascertaining the adequacy of each item in representing its hypothesised factor. Reporting the individual item factor loadings could have enhanced the clarity of this analysis. The respective PCA summaries outlined by Clough and colleagues also clearly show that a number of items were related to factors other than their hypothesised factor as demonstrated by cross-loadings. However, one cannot assess the extent to which these items cross-loaded as this information was not reported.

Table 2.2. Analysis summary of the 66-item measure of MT following PCA.

Factor	Variance accounted for (%)	No. of items retained (> 0.3)	Factor loading range	Cross loading factor	Decision made	No. of items removed from further analysis (< 0.3)
Commitment	8.5	11	0.31 to 0.69	Control – emotion item; “When I am feeling tired I find it difficult to get going”.	Item retained on control - emotion	0
Challenge	7.1	8	0.33 to 0.62	Commitment item; “I don’t usually give up under pressure”.	Item retained on commitment	3
Control - emotion	6.7	7	0.37 to 0.70	-	N/A	4
Control - life	6.0	7	0.41 to 0.61	Challenge item; “I often wish my life was more predictable”.	Item retained on challenge	4
Confidence - interpersonal	5.4	6	0.43 to 0.69	Control – emotion items; “When I am upset or annoyed I usually let others know” and “I generally hide my emotion from other”.	Both items retained on control – emotion (face validity)	0
Confidence – ability	5.2	9	0.39 to 0.60	-	N/A	2

Adapted from Earle (2006).

Table 2.3. Analysis summary of the MTQ48 six factor solution following PCA.

Factor	Variance accounted for (%)	No. of items retained (> 0.3)	Factor loading range	Cross loading factor	Decision made	No. of items removed from further analysis (< 0.3)
Challenge	15.1	8	0.34 to 0.63	Commitment item; "I don't usually give up under pressure".	Item retained on commitment	3
Control - life	13.5	7	0.41 to 0.61	Challenge item; "I often wish my life was more predictable".	Item retained on challenge	4
Commitment	11.3	11	0.32 to 0.69	Control – emotion item; "When I am feeling tired I find it difficult to get going".	Item retained on control - emotion	0
Confidence - abilities	9.3	9	0.40 to 0.61	-	N/A	0
Control - emotion	7.4	7	0.38 to 0.70	-	N/A	4
Confidence - interpersonal	6.1	6	0.43 to 0.69	Control - emotion item; "When I am upset or annoyed I usually let others know" and control - emotion item; "I generally hide my emotion from others".	Item retained on commitment (face validity)	0

Adapted from Clough *et al.* (2007).

Inspection of Earle's (2006) PCA indicates that only 13 out of 66 items were removed due to factor loadings being below the cut-off of 0.3. The amount of items removed by insufficient factor loadings is therefore inconsistent with Earle's extracted six factor 48-item measure of MT. Given that no information is presented regarding this inconsistency, one cannot determine the procedures used to remove the remaining five items. In a similar vein, Clough *et al.* (2007) state that 11 items were removed from further analysis following their PCA. However, no information is provided to detail the nature of the further analyses conducted or how items were removed given that Clough *et al.* extracted a six factor 48-item solution to underpin the MTQ48. There are also concerns with the samples used to develop the six factor model of the MTQ48. For instance, one may question the participants used to develop the MTQ48. Despite Clough *et al.* (2002) stating that the MTQ48 is a valid measure of MT in sport, close inspection of the participants used by Clough *et al.* and Earle reveals that only 52 and 166 athletes were used to develop the MTQ48, respectively. Consequently, the limited number of athletes used to develop the MTQ48 may cast doubt surrounding its utility in sport.

One may also suggest that the PCA procedures used by Clough *et al.* (2007) and Earle (2006) could have been improved. Clough and colleague's use of a liberal 0.3 cut-off level when extracting the items to represent their six factor model are not consistent with the scale development literature in that Garson (2006) has suggested using a cut-off of 0.4 and Fornell and Larcker (1981) suggest using a cut-off of 0.5. Clough and colleague's criteria may have resulted in the acceptance of weak items which in turn, may have led to the acceptance of a less representative instrument to measure the 4/6C's model of MT. In addition, one may suggest that using multiple extraction methods could have enhanced the PCA procedures adopted. Specifically, one may question Clough and colleague's solitary use of eigenvalues as a means of identifying factors given that previous research has found this method to over- or underestimate the number of factors in the data set (e.g., Bandalos & Boehm-Kaufman, 2008; Jackson, 1993; Zwick & Velicer, 1986). It could be argued that Clough and colleagues could have used Cattell's (1966) scree test and Parallel Analysis (PA: Franklin, Gibson, Robertson, Pohlmann, & Fralish, 1995) to supplement the information gleaned from inspecting eigenvalues which in turn, could have resulted in a more accurate determination of the number of factors extracted in their PCA. Specifically, Cattell's scree test requires the researcher to visually inspect eigenvalues when plotted against the factors in the data set. Inspection of the 'point of inflexion' is used to determine the number of factors to extract in the data set (Tabachnick & Fidell, 2001). In PA, the eigenvalues derived from the research data are compared to those from a random sample matrix of identical dimensionality to the research data set (i.e., identical sample size). Component PCA eigenvalues which are greater than their respective component PA eigenvalues derived from the random data set are retained (Franklin *et al.*, 1995).

Upon evaluating the validation protocols administered by Clough *et al.* (2007) and Earle (2006), it is apparent that there was a reliance on PCA to examine the factor structure of the MTQ48. Given that EFA techniques (e.g., PCA) have been argued to be more appropriate than CFA in the early stages of scale development where there is little understanding of how the data will interact to form distinct components (Kelloway, 1995), Earle's initial use of PCA was provided justified. However, analytical techniques such as CFA which test the robustness of an instruments factor structure would have enhanced model testing and are required to statistically support the hypothesised model (e.g., MacKenzie *et al.*, 2011; Marsh, 2007). CFA (unlike EFA) is underpinned by a strong theoretical foundation that enables the researcher to specify a factor model in advance and subsequently force items to load on specific factors (e.g., Jöreskog & Sorbom, 1993, 1996; Schutz & Gessaroli, 1993). Many researchers (e.g., Hagger & Chatzisarantis, 2009; MacKenzie *et al*; Marsh, 1997, 2002, 2007) have advocated the use of state-of-the-art analytical approaches such as CFA to test the factorial validity of psychometric instruments. Consequently, it could be argued that the preliminary research conducted by Clough and colleagues did not comprehensively examine the factorial validity of the MTQ48.

Research using the MTQ48 (e.g., Crust & Swann, 2011a, 2011b) has primarily used the research by Horsburgh *et al.* (2009) to support its factorial validity. Using a sample of the general population (n = 438), Horsburgh *et al.* tested the factor structure of the MTQ48. EFA and CFA findings showed superior model fit for the first-order four factor solution when compared to the unidimensional single factor solution. However, there are a number of limitations regarding the application of these analyses. The most pronounced and discussed limitation is the absence of any empirical data (i.e., fit indices, parameter estimates) to support their conclusion about the superiority of the first-order four factor solution when compared with a unidimensional, single factor solution (Gucciardi *et al.*, 2012). The absence of the relevant EFA and CFA statistics therefore makes it difficult for the reader to draw any robust conclusions regarding the adequacy of the first-order four factor model to underpin the MTQ48. In addition, Horsburgh *et al.'s* analyses would have been strengthened by testing the hypothesised six factor model and the second-order representations of the four and six factor models of the MTQ48. Given that Clough *et al.* (2007) and Earle (2006) extracted a six factor model to underpin the MTQ48, Horsburgh *et al.'s* analyses would have been enhanced by using CFA to test its adequacy in measuring the 4/6C's model of MT.

Moreover, the primary purpose of Horsburgh *et al.'s* (2009) study was not to examine the factorial validity of the MTQ48. Indeed, it was to examine the extent to which genes and/or environment factors contribute to the development of individual differences in MT and to determine the genetic and/or environmental basis of any relationship between MT and personality. This appears to have been reflected in data reported by Horsburgh *et al.* (2009) to support the MTQ48's factorial validity. A further limitation to Horsburgh *et al.'s* research was that no athletes were used in their factorial examination of the MTQ48 in that participants represented the general population. As previously discussed with the validation protocols administered by Earle (2006) and Clough *et al.* (2007), the lack of athletes used by Horsburgh *et al.* appears to be a distinct limitation in the efforts to validate the MTQ48 as a measure of MT in sport. This issue is further compounded given that previous inventories in the MT literature (e.g., Gucciardi, 2009) and beyond (e.g., Lane, Harwood, Terry, & Karageorghis, 2004; Martens & Webber, 2002) have failed to generalise to other samples or contexts other than those at which they were originally developed with.

At present, only one published study has used CFA to examine the factorial validity of the MTQ48. Gucciardi *et al.* (2012) examined the factorial validity of the MTQ48 in two independent samples; a) athletes, and b) full-time employees. Gucciardi *et al.* emphasised the importance of validating the MTQ48 within a workplace context as well as an athletic population given that the conception of the 4C's model of MT stems from hardiness theory. Participants were 686 athletes and 639 full-time employees who completed the MTQ48 via an online survey. CFA using the Robust Maximum Likelihood estimation method revealed that model fit for the first-order four factor solution was unsatisfactory for the athlete (x^2 [1074] = 5511.88, $p < 0.001$, Comparative Fit Index [CFI] = 0.487, Tucker-Lewis Index [TLI] = 0.462, Standardised Root Mean Square Residual [SRMR] = 0.104, Root Mean Square Error of Approximation [RMSEA] = 0.078, 90% confidence interval [0.076, 0.080]) and workplace sample (x^2 [1074] = 4928.95, $p < 0.001$, CFI = 0.521, TLI = 0.497, SRMR = 0.093, RMSEA = 0.075, 90% confidence interval [0.073, 0.077]). In addition to the poor model fit, the first-order four factor solution of the athlete and workplace sample was improper, as reflected by a factor correlation between the control and confidence factors that exceeded 1.0 (Blunch, 2008). Model fit indices and parameter estimates yielded by CFA did not support the hypothesised first-order four factor model of the MTQ48 with an athlete and workplace sample. The preliminary findings yielded by Gucciardi *et al.* therefore provide little support for the factorial validity of the MTQ48.

However, there are a number of limitations of Gucciardi *et al.'s* (2012) research which may reduce our confidence in their findings. First, Gucciardi *et al.* did not report a factorial analysis of the hypothesised six factor model of the MTQ48. Consequently, no research has rigorously used CFA to examine the factorial validity of the six factor model of the MTQ48. Second, Gucciardi *et al.* did not report factorial analyses of the second-order representations of the MTQ48. Given that research by Horsburgh *et al.* (2009) and Gucciardi *et al.* have only examined the first-order representations of the MTQ48, there is no information in the literature as to the adequacy of the second-order representations of the four and six factor models of the MTQ48.

Finally, Gucciardi *et al.* (2012) used an online survey to collect data. Although research by Lonsdale, Hodge, and Rose (2006) has provided support (e.g., response rates, missing data, coding errors) for the equivalency of online survey methods when compared to paper forms, research outside of sport and exercise psychology has been inconsistent (e.g., Birnbaum, 2004; Wygant & Lindorf, 1999). Consequently, research is required to test the factorial validity of the MTQ48 when using paper forms of the measure. Notwithstanding, the findings of Gucciardi *et al.* (2012) indicate little support for the factorial validity of the MTQ48. Although construct validation is an ongoing process which requires replication and verification across different samples of the target population (Marsh, 1997, 2002, 2007), the preliminary evidence yielded by Gucciardi *et al.* does suggest that there are problems with the MTQ48. Consequently, further research is required to comprehensively examine the factorial validity of the MTQ48 to determine its utility in measuring the 4/6C's model of MT in sport.

2.4.2.1 MacKenzie et al.'s (2011) Scale Development Framework

The concerns relating to the MTQ48's factorial validity (e.g., Connaughton & Hanton, 2009; Gucciardi *et al.*, 2011) largely stem from the limited research which has used CFA to examine its underpinning factor structures. Although research is required to re-examine the MTQ48's factorial validity, research which does not provide support may be due to inadequacies in the scale development procedures used. Despite there being a number of scale development guidelines (e.g., Churchill, 1979; Clark & Watson, 1995; Hinkin, 1995; Kline, 2000; MacKenzie 2003; Sartori, 1984) available to Earle (2006), MacKenzie *et al.'s* (2011) Overview of the Scale Development procedure (see Figure 2.1) repackages these guidelines into an operational framework. Consequently, when looking through the lens of MacKenzie and colleagues, one can evaluate the development and validation of the MTQ48.

2.4.2.2 MacKenzie et al.'s (2011) Step 1

The first important aspect of the scale development process requires the researcher to develop a conceptual definition of the construct (MacKenzie, 2003). Figure 2.2 provides a pictorial representation of the collective recommended guidelines outlined in the scale development literature for defining focal constructs. According to MacKenzie (2003), failure to adequately define the conceptual domain of a construct leads to a number of problems. First, it leads to confusion as to what the construct refers to, and its subsequent relations and differences with other constructs that already exist. Second, it leads to the development of poor items because the definition of the target construct is not adequately representing its underpinning literature and/or because the definition overlaps with existing constructs. Finally, it can lead to invalid conclusions about relationships with other constructs that later have to be rejected on the grounds that the items of the target construct are not adequately representing what they intend to capture.

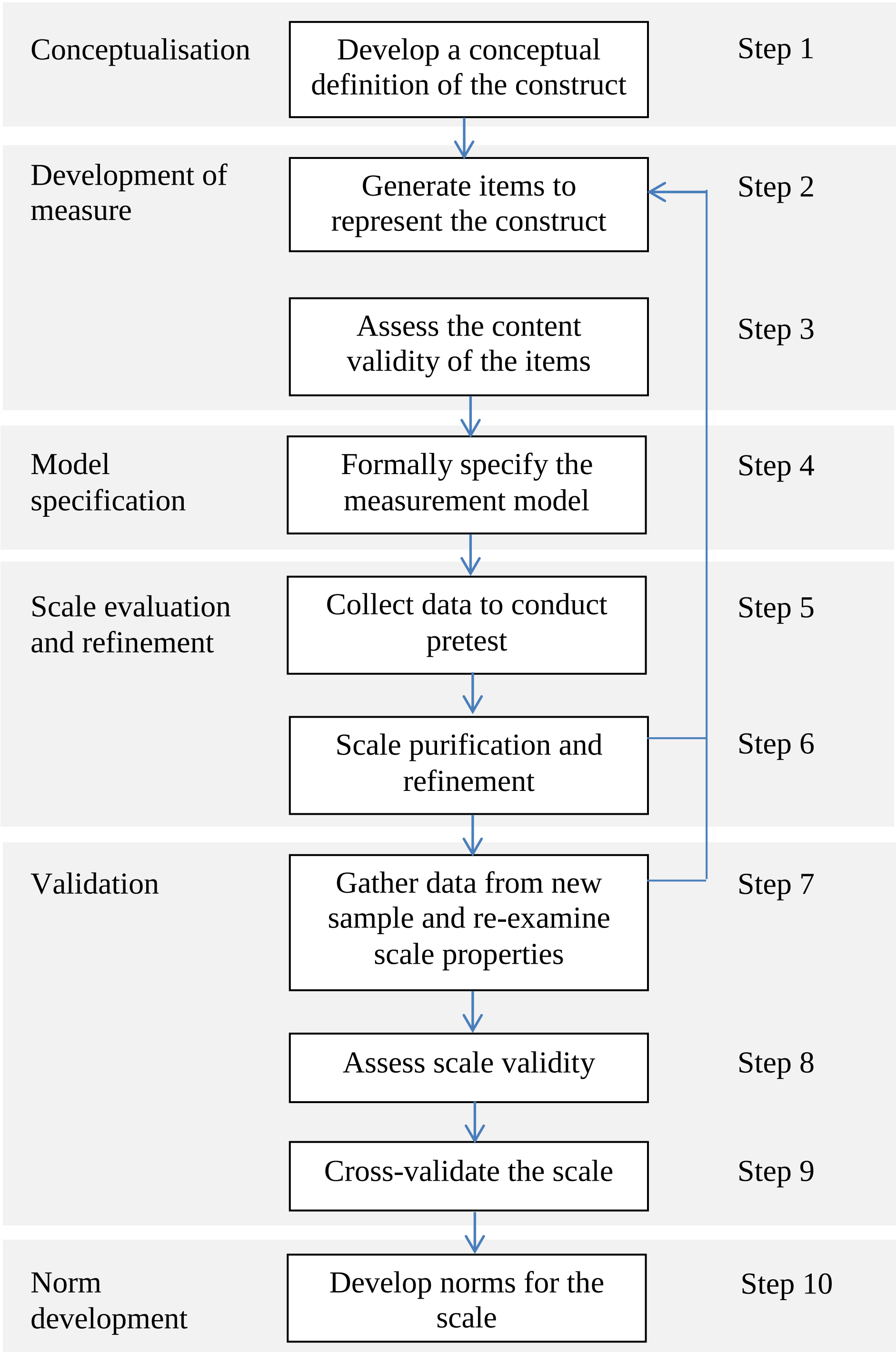


Figure 2.1. MacKenzie *et al.'s* (2011) Overview of Scale Development Procedure.

Step 1	
Principle: Articulate the basic construct	Method: Review the relevant literature to which will serve to clarify the nature and range of the content of the target construct (Clark & Watson, 1995).

Step 2	
Principle: Construct deduction	Method: Collect a representative set of definitions, extract their characteristics, and construct matrices that organise such characteristics meaningfully (Sartori, 1984).

Step 3	
Principle: Specify the nature of the construct	Method: Specify the conceptual domain to which the focal construct belongs, and the entity to which it applies in unambiguous terms and in a manner that is consistent with prior research. By conceptual domain, the definition should specify whether the construct refers to a thought, a feeling, a perception, an action, an outcome, or an intrinsic characteristic. By entity, the definition should specify the object to which the property applies, such a person, a task, a process, a relationship (MacKenzie *et al.*, 2011).

Step 4	
Principle: Generate construct definitions	Method: Consider steps 1 – 3 to inform the generation of a clear and concise construct definition (e.g., Churchill, 1979; Hinkin, 1995; MacKenzie *et al.*).

Figure 2.2. Pictorial amalgamation of the construct definition process.

2.4.2.3 Step 1 and the MTQ48

Although Clough *et al.* (2002) and Earle (2006) did not initially provide explicit definitions for the constructs underpinning the 4/6C's model of MT, this was achieved by Clough *et al.* (2007; see Figure 2.1). Close inspection of the content and structure of the information reported by Clough *et al.* appears to be inconsistent with the recommendations outlined in the scale development literature (see Figure 2.2). First, researchers should review the relevant literature to ensure that the content of the target construct is fully captured (Clark & Watson, 1995). Given that the 4/6C's model is underpinned by hardiness, Clark and Watson's guidelines would suggest that the challenge, commitment, and control constructs of the 4/6C's model of MT should be reflective of those characteristics. However, inspection of the characteristics articulated by Clough *et al.* appear to be partially different to those outlined in the hardiness literature. For instance, despite being successful in encapsulating the willingness of high challenge individuals to seek out challenges, Clough *et al.* do not capture information stating that change is viewed as a normal part of life (e.g., Maddi, Propst, & Feldinger, 1965). Furthermore, Clough *et al.'s* commitment characteristics do not adequately capture the characteristics outlined in the hardiness literature, including the notion that highly committed individuals find an interest in whatever they do (Maddi & Hightower, 1989).

Furthermore, one may question the content of the information reported by Clough *et al.* (2007) to underpin the confidence construct. Using Clark and Watson's (1995) guidelines, it could be argued that Clough *et al.* should have reviewed the dispositional confidence literature to inform the development of a theoretically embedded confidence definition. Although the confidence literature may have provided little guidance given the paucity of research which has examined dispositional confidence, Clough *et al.* could have utilised research from related psychological concepts to fulfil the recommendations of Clark and Watson. Specifically, the trait component of self-esteem (i.e., self-worth, self-competence) could have been used to identify the underlying characteristics of dispositional confidence which in turn, may have increased the likelihood of fully capturing its target content. There also appears to be problems with the structure of the information presented by Clough *et al.* in that characteristics are presented for the overarching control and confidence factors as well as their respective nested components.

Once the identified characteristics have been organised (Sartori, 1984), the researcher is required to formally specify the conceptual domain and entity of the construct (MacKenzie *et al.*, 2011; see Figure 2.2). Inspection of the information reported by Clough *et al.* (2007) reveals a lack of information relating to the conceptual domain and entity of the respective factors of the 4/6C's model of MT. The final stage of the construct definition process requires the researcher to consider steps 1-3 to inform the generation of a clear and concise construct definition (e.g., Churchill, 1979; Hinkin, 1995; MacKenzie *et al.*). Given the lack of explicit information relating to the conceptual domain and entity of the constructs underpinning the 4/6C's model of MT and the apparent emphasis on construct descriptors, it appears that Clough *et al.* (2007) did not generate clear and concise definitions and thus did not satisfy the final stages of the construct definition process. Definitions outlined in the hardiness literature can be used to qualify the apparent inadequacies of the information reported by Clough *et al*. For instance, Kobasa *et al.* (1982) defined control as the "control disposition is expressed as a tendency to feel and act as if one is influential (rather than helpless) in the face of the varied contingencies of life" (p. 169). Kobasa *et al.'s* definition clearly specifies the conceptual domain (i.e., disposition) and the entity (i.e., feel and act) of the construct and provides a clear and concise operational definition of control. Upon evaluating the content and structure of the definitions underpinning the 4/6C's model of MT, it appears that Clough *et al.* did not adhere to the available scale development guidelines (e.g., Churchill, 1979; Clark & Watson, 1995; Hinkin, 1995; Kline, 2000; MacKenzie 2003; Sartori, 1984) when constructing definitions to underpin the 4/6C's model of MT and ultimately the MTQ48.

2.4.2.4 MacKenzie et al.'s Step 2

The second important aspect of the scale development process requires the researcher to generate items to represent the construct. The ultimate goal of the item generation process is to produce a pool of items that fully captures the varying facets of the target construct while minimising the extent to which items tap into concepts other than the one under examination (e.g., MacKenzie, 2003; MacKenzie *et al.,* 2011). Researchers suggest that a variety of sources should be consulted when generating items including reviews of the literature, deduction from the theoretical definition of the construct (i.e., hardiness and confidence), suggestions from experts in the field, interviews or focus group discussions with representatives of the population(s) to which the focal construct is expected to generalise, and an examination of existing measures relating to the target construct (e.g., Clark & Watson, 1995; MacKenzie *et al.*). Additionally, it has been suggested that the wording of all items should be considered and analysed to ensure that item structure is as simple and precise as possible (e.g., Peterson, 2000; Podsakoff *et al.,* 2003; Spector, 1992; Torangeau, Rips, & Rasinski, 2000). In order to assess item clarity, a number of areas should be considered including the identification and splitting of double barrelled items, items possessing complicated syntax should be clarified and simplified, and items containing obvious social desirability should be refined or removed (e.g., Clark & Watson, 1995; MacKenzie *et al.*).

2.4.2.5 Step 2 and the MTQ48

Earle (2006) used a panel of experts to generate an initial pool of 66 items. Earle used data from the interviews conducted (see Chapter 2.2) to develop the 4/6C's model of MT and existing questionnaires were used to generate items which fully represented the breadth of the constructs underpinning the 4/6C's model of MT. Each item was written in accordance with the guidelines proposed by Kline (1986) to ensure that all items were understandable, unequivocal, and specific. Despite Earle using a panel of experts to generate items, greater information regarding their expertise (e.g., psychometrics, hardiness, MT) would have strengthened the readers understanding of the procedures undertaken. Using the scale development literature as a guiding framework (e.g., Clark & Watson, 1995; MacKenzie *et al.*, 2011), there appears to be concerns regarding the content and structure of the MTQ48 items. In an effort to explore the MTQ48 items, a cursory assessment of item content and structure was undertaken. Specifically, the author of this thesis assessed the content and structure of each item against the definitions forwarded by Clough *et al.* (2007; see Table 2.1) to evaluate its representation. According to Gucciardi *et al.* (2011), face validity is “the extent to which a measure appears (i.e., “on the surface”) to assess what it claims to capture” (p. 111). Table 2.4 displays the exploration of MTQ48 items and may indicate their adequacy to measure the respective factors of the 4/6C’s model of MT. If the item appeared to match a characteristic forwarded to make-up its hypothesised factor, it may indicate good face validity (depicted by “Yes” followed by its matching characteristic). If the item did not appear to match a characteristic forwarded by Clough and colleagues, it may indicate poor face validity (depicted by “No” followed by “does not match Clough *et al.’s* definition”). Exploration of MTQ48 item face validity suggests that only 30 out of 48 items appear to adequately represent their hypothesised factor definitions (see Table 2.1 for Clough *et al.’s* 4/6C’s definitions).

Table 2.4. Exploration of MTQ48 item representation relative to Clough *et al.'s* (2007) factor definitions.

Factor		Item	Representation
Challenge	MTQ4	Challenges usually bring out the best in me	Yes; "individuals see challenges as opportunities"
	MTQ6R	Unexpected changes to my schedule generally throw me	Yes; "minimise their exposure to change and the problems that come with that"
	MTQ14R	I often wish my life was more predictable	No; does not match Clough *et al.'s* definition
	MTQ23	I generally cope well with any problems that occur	No; does not match Clough *et al.'s* definition
	MTQ30	I am generally able to react quickly when something unexpected happens	No; does not match Clough *et al.'s* definition
	MTQ40	I usually look forward to changes in my routine	Yes; "Individuals who see them as opportunities will actively seek them out"
	MTQ44	I usually enjoy a challenge	Yes; "Individuals who see them as opportunities will actively seek them out"
	MTQ48	I can usually adapt myself to challenges that come my way	No; does not match Clough *et al.'s* definition
Commitment	MTQ1	I usually find something to motivate me	No; does not match Clough *et al.'s* definition
	MTQ7	I don't usually give up under pressure	Yes; "the ability for an individual to carry out tasks successfully despite any problems or obstacles"
	MTQ11R	"I just don't know where to begin" is a feeling I usually have when presented with several things to do at once	No; does not match Clough *et al.'s* definition
	MTQ19	I can generally be relied upon to complete the tasks I am given	Yes; "the ability for an individual to carry out tasks successfully despite any problems or obstacles"
	MTQ22R	I am easily distracted from tasks that I am involved with	No; does not match Clough *et al.'s* definition
	MTQ25	I generally try to give 100%	No; does not match Clough *et al.'s* definition
	MTQ29R	When faced with difficulties I usually give up	Yes; "the ability for an individual to carry out tasks successfully despite any problems or obstacles"
	MTQ35R	I usually find it difficult to make a mental effort when I am tired	No; does not match Clough *et al.'s* definition

Table 2.4. *(continued).*

Factor		Item	Representation
Commitment *cont.*	MTQ39	I can normally sustain high levels of mental effort for long periods	No; does not match Clough *et al.'s* definition
	MTQ42R	I usually find it hard to summon enthusiasm for the tasks I have to do	No; does not match Clough *et al.'s* definition
	MTQ47R	When I face setbacks I am often unable to persist with my goal	Yes; "the ability for an individual to carry out tasks successfully despite any problems or obstacles"
Control - emotion	MTQ21R	I generally find it hard to relax	Yes; "able to keep anxieties in check"
	MTQ26R	When I am upset or annoyed I usually let others know	Yes; "are less likely to reveal their emotional state to other people"
	MTQ27R	I tend to worry about things well before they actually happen	Yes; "able to keep anxieties in check"
	MTQ31	Even when under considerable pressure I usually remain calm	Yes; "able to keep anxieties in check"
	MTQ34	I generally hide my emotion from others	Yes; "are less likely to reveal their emotional state to other people"
	MTQ37R	When I am feeling tired I find it difficult to get going	No; does not match Clough *et al.'s* definition
	MTQ45	I can usually control my nervousness	Yes; "able to keep anxieties in check"
Control - life	MTQ2	I generally feel in control	Yes; "more likely to believe that they control their lives"
	MTQ5	When working with other people I am usually quite influential	Yes; in relation to Clough *et al.'s* overarching control factor – "capable of exerting more influence on their working environment"
	MTQ9R	I usually find myself just going through the motions	No; does not match Clough *et al.'s* definition
	MTQ12	I generally feel that I am in control of what happens in my life	Yes; "more likely to believe that they control their lives"
	MTQ15R	Whenever I try to plan something, unforeseen factors usually seem to wreck it	Yes; "They feel that their plans will not be thwarted and that they can make a difference"
	MTQ33R	Things just usually happen to me	No; does not match Clough *et al.'s* definition
	MTQ41R	I feel that what I do tends to make no difference	No; does not match Clough *et al.'s* definition

Table 2.4. *(continued).*

Factor		Item	Representation
Confidence - ability	MTQ3	I generally feel that I am a worthwhile person	Yes; “more likely to believe that they are a truly worthwhile person”
	MTQ8	I am generally confident in my own abilities	Yes; in relation to Clough *et al.’s* overarching confidence factor – “Individuals who are high in confidence have the self-belief to successfully complete tasks”
	MTQ10R	At times I expect things to go wrong	No; does not match Clough *et al.’s* definition
	MTQ13	However bad things are, I usually feel they will work out positively in the end	Yes; “tend to be more optimistic about life in general”
	MTQ16	I generally look on the bright side of life	Yes; “tend to be more optimistic about life in general”
	MTQ18R	At times I feel completely useless	Yes; “more likely to believe that they are a truly worthwhile person”
	MTQ24	I do not usually criticise myself even when things go wrong	Yes; in relation to Clough *et al.’s* overarching confidence factor – “individuals at one end of the scale will be able to take setbacks (externally and self-generated) in their stride”
	MTQ32R	If something can go wrong, it usually will	No; does not match Clough *et al.’s* definition
	MTQ36R	When I make mistakes I usually let it worry me for days after	Yes; in relation to Clough *et al.’s* overarching confidence factor – “At the other end individuals will be unsettled by setbacks and will feel undermined by these”
Confidence - interpersonal	MTQ17	I usually speak my mind when I have something to say	Yes; “They are less likely to be intimidated in social settings”
	MTQ20	I usually take charge of a situation when I feel it is appropriate	Yes; “more likely to push themselves forward in groups”
	MTQ28R	I often feel intimidated in social gatherings	Yes; “They are less likely to be intimidated in social settings”
	MTQ38	I am comfortable telling people what to do	Yes; “more likely to push themselves forward in groups”
	MTQ43	If I feel somebody is wrong, I am not afraid to argue with them	No; does not match Clough *et al.’s* definition
	MTQ46R	In discussions, I tend to back-down even when I feel strongly about something	Yes; “They are less likely to be intimidated in social settings”

Note: Clough *et al.’s* (2007) factor definitions are displayed in Table 2.1.

Specifically, 4 out of 8, 7 out of 11, 1 out of 7, 3 out of 7, 2 out of 9, and 1 out of 6 items appear to be poor representations of Clough *et al.'s* (2007) definitions of challenge, commitment, control – emotion, control – life, confidence – abilities, and confidence – interpersonal, respectively. A number of items also appear to represent the overarching characteristics of their hypothesised factor as opposed to the characteristics relating to their hypothesised nested factor. For instance, item MTQ5 (confidence – life) appears to represent the characteristics relating to the overarching control factor and items MTQ8, MTQ24, and MTQ36R (confidence – abilities) appear to represent the characteristics relating to the overarching confidence factor. Research by Gucciardi *et al.* (2012) has also questioned the content of the MTQ48 items by stating that item MTQ3 and item MTQ18R appear to be capturing aspects of self-esteem and item MTQ13 and item MTQ16 appear to be capturing one's perceived levels of optimism.

In addition to item content, a number of items appear to be inadequate with regards to their structure. For example, item MTQ11R is the only item which incorporates the use of quotations which is inconsistent with the remaining MTQ48 items (see Table 2.4). Moreover, item MTQ26R uses double barrelled wording (see Table 2.4) which conflicts with the recommendations of the scale development literature (e.g., Clark & Watson, 1995; MacKenzie *et al.,* 2011). Inspection of MTQ48 items reveals that a number of items (e.g., MTQ9R, MTQ11R, MTQ16, MTQ30R, MTQ33R, MTQ37R, MTQ41R) are quite ambiguous and lack specificity which may therefore reduce the accuracy of responses to these items (see Table 2.4). Upon evaluating the content and structure of the MTQ48 items relative to the scale development literature, one may not only question the MTQ48's ability to adequately represent the 4/6C's model of MT, but also the interpretability of its items. The evidenced inadequacies of the MTQ48 items suggest that the item generation procedures adopted by Earle (2006) could have been strengthened. Thus, this may be an issue contributing towards current failures to validate the measure using CFA.

2.4.2.6 MacKenzie et al.'s (2011) Step 3

The third important aspect of the scale development process requires the researcher to evaluate items for their content validity. According to Gucciardi *et al.* (2011), content validity is "the degree to which the items of a measure sufficiently represent all facets (i.e., entire domain) of the construct of interest" (p. 111). Although MacKenzie *et al.* (2011) offer a number of methods to evaluate item content validity, Lynn's (1986) CVI has recently been advocated as an effective method within the sport and exercise psychology literature (e.g., Bartholomew, Ntoumanis, & Thogersen-Ntoumani, 2010; Williams & Cumming, 2011). Specifically, Lynn's CVI requires independent scrutineers to evaluate the content of items relative to their hypothesised construct definitions and is calculated by dividing the number of scrutineers who give a rating of 4 or 5 (i.e., rate items to be a good match or excellent match to the hypothesised definition, respectively) by the number of scrutineers. Lynn (1986) suggests that when a minimum of 6 scrutineers are used CVI's close to 0.80 are acceptable. Item CVI's are then used to inform item acceptance, refinement, and removal.

2.4.2.7 Step 3 and the MTQ48

Earle (2006) used 20 athletes to pilot the initial item pool to ensure that each item was unambiguous and clear. Findings revealed that all items were thought to be suitable and understood. Despite Earle indicating support for the representation of the MTQ48 items, the previously discussed concerns relating to the MTQ48's item content and structure may question this conclusion (see Chapter 2.4.2.5). Moreover, the procedures used could have been improved by using an established method to assess content validity (e.g., CVI, Lynn, 1986). Using Lynn's recommendations as a guide, it could be argued that Earle could have used this phase of the scale development process to remedy the apparent deficiencies of the MTQ48 items. Indeed, this could have led to the acceptance, refinement, and removal of items which in turn, could have resulted in a pool of items which not only structurally conform to the guidelines outlined in the scale development literature, but also more accurately represent their hypothesised factors. The apparent deficiencies with the MTQ48 items coupled with the absence of an established method to evaluate item content validity suggests that Earle could have strengthened the procedures used to fulfil this aspect of the scale development process.

2.4.2.8 MacKenzie et al.'s Step 4

The fourth important aspect of the scale development process requires the researcher to formally specify the measurement model. The expected relationships between the items and their hypothesised factors and/or subcomponents they are intended to capture should be specified (e.g., MacKenzie, 2003; Mackenzie *et al.*, 2011). According to Marsh (2002; 2007), the CFA measurement model requires the researcher to specify an *a priori* factor structure in advance of examining the relations among the latent constructs. In this highly restrictive approach, no cross-loadings are allowed and all nontarget loadings are constrained to zero (Thompson, 2004). CFA enables the researcher to evaluate the validity of individual items by determining whether the relationship between each item and its hypothesised latent construct is large and significant (e.g., Jöreskog & Sorbom, 1993, 1996; Schutz & Gessaroli, 1993). The measurement model therefore specifies the latent constructs in a heterarchical arrangement and assumes that all latent constructs are equal. When constructing the measurement model, researchers are also required to specify whether the constructs in the model are best represented in a first- or second-order arrangement (Marsh, 2002; 2007). According to Chen, Sousa, and West (2005), first-order models are arranged so that each latent construct is specified so that it correlates with the remaining latent constructs, whereas second-order models are arranged so that the seemingly distinct, yet related latent constructs are accounted for by a higher order construct(s).

Conversely, if some latent constructs in the model are theoretically predicted by other latent constructs in the model, structural models have been offered as a more appropriate means to test factorial validity (Marsh, 2007). The structural model therefore specifies the latent constructs in a hierarchical arrangement and assumes that some latent constructs will predict others. Despite there being differences between how the measurement model and structural model are arranged, they share one common principle – the specification of an *a priori* model structure.

2.4.2.9 Step 4 and the MTQ48

As previously discussed, Clough *et al.* (2007) and Earle (2006) used PCA as a means to develop and validate the MTQ48. Given that PCA does not require an *a prior* model structure, Clough and colleagues provided no information to formally specify the factor structure of the MTQ48. Subsequently, this step of the scale development process was overlooked. Information relating to the MTQ48's model specification is particularly salient given the presence of the competing four (Clough *et al.,* 2002) and six (Earle) factor model structures purporting to underpin the instrument and the possible alternative arrangements of these models (i.e., first- vs. second-orders models). Specifically, the absence of CFA to examine the MTQ48's factorial validity when arranged in a second-order model could provide a particularly fruitful avenue of investigation for future research. Inspection of the theoretical propositions of hardiness indicates that Kobasa (1979) conceptualised hardiness as a single, unitary construct which comprises of three sub-ordinate dimensions of challenge, commitment, and control. In other words, "it is the combination of all 3C's that constitutes hardiness" (Maddi, 2002, p. 176). Previous instruments measuring hardiness have provided support for its hypothesised second-order representation whereby the first-order factors combine to constitute a broader hardiness dimension (e.g., Revised Dispositional Resilience Scale-15 (DRS15-R); Bartone, 1995). In light of the 4/6C's model of MT being underpinned by hardiness, one could argue that the MTQ48 could be best represented by a second-order model whereby the constructs of challenge, commitment, control (control – emotion, control life), and confidence (confidence –abilities, confidence – interpersonal) are accounted for by a higher order MT construct.

However, Carver (1989) questioned this approach and advocated examining the separate effects of the hardiness subcomponents. According to Carver, combining conceptually related yet distinct constructs may result in substantial information loss when interpreting how hardiness relates to other constructs. When constructs are integrated, researchers lose the ability to assess their relative importance and their respective and interactive effects on dependent measures. For example, research examining the effects of hardiness upon stress reactions in military contexts revealed that the challenge construct has been found to provide more information than the general hardiness factor (Eid, Johnsen, Saus, & Risberg, 2004; Eid & Morgan, 2006). As a result, recent research (e.g., Ford *et al.*, 2000; Klag & Bradely, 2004) has tended to examine the effect of both the total hardiness score and each of its subcomponents. In line with this perspective, it could be argued that the MTQ48 is best represented by a first-order model whereby each latent construct is specified to be correlated with the other constructs in the model. Given that previous research (e.g., Gucciardi *et al.,* 2012; Horsburgh *et al.,* 2009) has only examined the factorial validity of the first-order four factor model of the MTQ48 and that researchers have emphasised the need to examine the dimensionality of an instrument's conceptual model (e.g., Gignac, 2009; Marsh, Martin, & Jackson, 2010), future research is required to examine all possible model structures underpinning the MTQ48 (i.e., four vs. six, first- vs. second-orders models).

Additionally, previous research has not considered examining the MTQ48's factorial validity when using structural models. Specifically, previous research (e.g. Gucciardi *et al.*, 2012; Horsburgh *et al.,* 2009) has used measurement models to conceptually represent the challenge, commitment, control, and confidence constructs in a heterarchical arrangement whereby all constructs are assumed to be equal. However, it may be conceptually plausible that some of the latent constructs in the 4/6C's model of MT predict other constructs in the model. For example, inspection of the propositions of Maddi (1990, 1997, 2002, 2004) may suggest that it is conceptually plausible for hardiness to predict confidence. According to Maddi, hardiness is embedded within existential theory (Frankl, 1959; Gendlin, 1966; Kierkegaard, 1954) and is purported to provide the existential courage that facilitates the individual to strike out and discover the future despite its uncertainty (Maddi, 2004; Tillich, 1952). If we assume that existential courage shares similar characteristics to constructs such as confidence, it is conceptually plausible for challenge, commitment, and control to predict confidence in the 4/6C's model of MT. Specifically, dispositional confidence may be a result of an individual's belief that change is a normal part of life (challenge), having a tendency to involve oneself in whatever one is doing (commitment), and having a tendency to feel and act as if one is influential (control).

In contrast, the theoretical predictions of challenge, commitment, and control may offer different propositions. For instance, the Theory of Challenge and Threat States in Athletes (Jones, Meijen, McCarthy, & Sheffields, 2009) proposes that self-efficacy is a determinant of a challenge/threat state. Bandura (1986) defined self-efficacy as a belief in one's capability or skill to attain a particular goal or execute a particular behaviour which can explain not only the choice and level at which an activity is pursued, but also the likelihood of successful completion of the behaviour. Jones *et al.* suggest that the belief that one has the necessary skills to execute the courses of action required to succeed clearly contributes to the perception that he/she has the resources to successfully cope with the demands of the situation. If we accept the similarity between self-efficacy and confidence, it is conceptually plausible for confidence to predict one's tendency to perceive new situations, changes, or experiences as challenging and stimulating (challenge), and one's tendency to believe and act as though they have the power to influence the outcome of a situation (control). Given the lack of understanding of the MTQ48's model specification, it is apparent that research is warranted to examine the possible model structures underpinning the instrument to determine its adequacy in measuring the 4/6C's model of MT.

2.4.2.10 MacKenzie et al.'s Step 5 *to 9*

Given that the remaining steps of the scale development process are primarily concerned with examining the psychometric properties of an instrument (once developed), the respective steps will be discussed collectively. According to Marsh (e.g., 1997, 2002, 2007), construct validation is central to the development and validation of psychometric tools. Marsh emphasises the need to continually evaluate an instrument's validity and reliability within a construct validation framework whereby its utility is judged upon theoretical, measurement, empirical research, and practical grounds. Specifically, Marsh emphasises the need to construct a sound theoretical framework (i.e., relations with other constructs in a theoretically justifiable manner) and definition of the target construct prior to conducting analytical techniques which serve to validate the conceptualisation. Once these areas have been satisfied, researchers typically examine the within- and between-network properties of the target construct (Marsh, 2002).

Within-network properties concerns the examination of an instrument's internal factor structure, its invariance across relevant subgroups of the target population (e.g., gender, age), and its item properties (e.g., distributional, descriptive; Marsh, 2002). Statistical techniques such as EFA, CFA, and reliability analyses which test an instrument's internal structure are commonly employed. Specifically, internal consistency and test-retest reliability are frequently used to assess an instrument's reliability. Internal consistency is based on intercorrelations among items to estimate response consistency, with Cronbach's Alpha being the most commonly used method (e.g., Vaughn, Lee, & Kamata, 2012). Test-retest reliability is concerned with the extent to which constructs remain stable over time (Lane, Nevill, Bowes, & Fox, 2005) and is considered vital in validating psychometric tools (e.g., Anastasi & Urbina, 1997; Kline, 1993). Researchers (e.g., McCrae, Kurtz, Yamagata, & Terracciano, 2011) have further emphasised the importance of obtaining stability of personality test scores where the target trait(s) is expected to be relatively stable over time. Although examination of both within- and between-network validity are important to the construct validity enterprise, researchers (e.g., Gignac, 2009; Marsh & Hau, 2007) have emphasised the importance of establishing within-network validity prior to assessing between-network validity.

Between-network properties concern examining the relationship of the target construct with salient demographic variables (e.g., age, playing experience), cognitions (e.g., self-determination, anxiety symptoms) and behaviours (e.g., performance; Marsh, 2002). Central to the assessment of between-network validity is the identification of the constructs' nomological network whereby the antecedents and consequences of the target construct are postulated (Cronbach & Meehl, 1955). Analytical techniques such as correlation, regression, and cluster analyses are frequently used to assess such relationships (Marsh). Due to the varying nature of the relationships examined in this strand of analysis, a number of validity types exist within the literature. Of particular importance for between-network validity is the examination of convergent, predictive, and criterion validity (Marsh).

Convergent validity is concerned with assessing theoretically hypothesised relationships between scores from one measure and those from another measure (Vaughn & Daniel, 2012). For instance, if we assume MT and coping effectiveness to be theoretically related, convergent validity would be supported by obtaining a positive correlation between the scores yielded to measure these constructs. Convergent validity is commonly assessed when both measures are collected at the same time (also known concurrent validity). Despite Marsh (2007) positing that large correlations (i.e., $r > 0.70$) are indicative of convergent validity when assessing the relations between instruments purporting to measure the same or substantially overlapping scales, the guidelines to inform convergent validity when assessing nomological validity appear to be less clear. According to Kline (2005), correlations ranging from 0.30 to 0.50 are indicative of nomological validity. However, recent research (e.g., Freeman, Coffee, & Rees, 2011; Lonsdale, Hodge, & Rose 2008; Williams & Cummings, 2011; Zourbanos, Hatzigeorgiadis, Chroni, Theodorakis, & Papaioannou, 2009) in the sport and exercise psychology literature has argued that nomological validity can be demonstrated by assessing correlation patterns. Specifically, these researchers have argued that if correlations obtained are significant and in the hypothesised direction, one can indicate convergent validity.

Predictive validity is an extension of convergent validity and is concerned with the extent to which scores from one measure predict scores from another measure which it is hypothesised to be related (Vaughn & Daniel, 2012). For example, if we assume MT and challenge appraisals to be theoretically related, we would expect MT to be positively related to challenge appraisals and for MT to predict future appraisals to stress. Although it is common for research to use cross-sectional data to assess predictive validity, Gignac (2009) argues that it is more impressive to collect responses to the predictor variable at one time point and the outcome variable at a later time point. Criterion validity is the extent to which a target construct(s) predicts a theoretically justified outcome or behaviour (e.g., performance; Vaughn & Daniel). For instance, if we assume a theoretical association between MT and perseverance on an endurance task, criterion validity would be supported by establishing differences in endurance performance based on MT scores. According to Marsh (2002), the evidence yielded from the respective components of construct validity enables the researcher to determine the utility of the instrument, and ultimately its underpinning conceptualisation.

2.4.2.11 Steps 5 to 9 and the MTQ48

Within-network properties

As previously discussed, insufficient attention has been given to examine the factorial validity of the MTQ48, with preliminary research (e.g., Gucciardi *et al.*, 2012) providing little support for its factor structure. Although research has not explicitly examined the MTQ48's factorial invariance, research examining differences in MTQ48 scores across a number of demographic variables (e.g., gender, playing standard) has been inconsistent (see Chapter 2.3.1: e.g., Crust & Azadi, 2010; Crust & Keegan, 2010; Nicholls *et al.*, 2009). With regards to test-retest reliability, Earle (2006) reported adequate reliability estimates for the MTQ48 using a sample of 108 psychology undergraduate students (*M* age = 19.22, *S.D.* = 2.81). Findings showed that the MTQ48 was stable over a six week interval in that Pearson's correlation coefficients were high for all subscales, ranging from 0.80 (challenge) to 0.87 (for control – emotion).

Many researchers (e.g., Crust & Clough, 2005; Crust & Swann, 2011a) have used the MTQ48's internal reliability estimates as a means of supporting its within-network properties. However, close inspection of the documented estimates reveals inconsistent findings. According to Nunnally and Bernstein (1994), Cronbach's Alpha reliability coefficients (α) greater than 0.70 are acceptable. Earle (2006) reported acceptable internal reliability values of total MT and the respective subscales of the MTQ48; challenge: $\alpha = 0.71$; commitment: $\alpha = 0.80$; control: $\alpha = 0.74$; control - emotion: $\alpha = 0.70$; control - life: $\alpha = 0.72$; confidence: $\alpha = 0.81$; confidence - abilities $\alpha = 0.75$; confidence - interpersonal $\alpha = 0.76$; total MT: $\alpha = 0.90$. Although further research has found support for the internal reliability of the MTQ48 for total MT scores (Crust, 2009; Crust & Keegan, 2010; Crust & Swann; Kaiseler *et al.*, 2009; Nicholls *et al.*, 2008, 2011), these studies have shown some subscales to be inadequate (i.e., $\alpha < 0.70$; Nunnally & Bernstein) including challenge (Crust & Keegan; Nicholls *et al.*, 2011), commitment (Crust & Keegan), control – emotion (Crust & Keegan; Crust & Swann; Kaiseler *et al.*, 2009), and control – life (Crust & Keegan; Crust & Swann). The inconsistent evidence relating to the MTQ48's internal reliability appears to add further concern regarding its ability to measure the 4/6C's model of MT.

Between-network properties

Despite the lack of evidence to support the factorial validity of the MTQ48 and the need to assess factorial validity prior to assessing between-network validity (Gignac, 2009; Marsh & Hau, 2007), Earle (2006) examined the convergent validity of the MTQ48. Using a sample of 106 undergraduate students (*M* age = 19.22, *S.D.* = 2.81), MTQ48 subscales scores were correlated with a range of personality measures. Findings revealed that the MTQ48 subscales were significantly related to scores on the Life Orientation Test (r range = 0.39 to 0.53), Satisfaction with Life Scale (r range = 0.49 to 0.59), Self-esteem Scale (r range = 0.34 to 0.49), Self-efficacy (r range = 0.59 to 0.74), and State Trait Anxiety Inventory (r range = -0.54 to -0.63). Clough *et al.* (2007) examined convergent validity by correlating scores with eight subscale scores of the PREVUE personality scale (Bartram, 1994). Findings showed that the MTQ48 was significantly related with the PREVUE dimension of "Excitable-Relaxed" (r = 0.48), "Restless-Poised" (r = 0.38), "Submissive-Assertive" (r = 0.38), "Reserved – Outgoing" (r = 0.39), and "Self-sufficient-Group orientated" (r = 0.24), but was not significantly related to "Co-operative-Competitive" (r = 0.20), "Reactive-Organised" (r = -0.05), and "Innovative-Conventional" (r = 0.07). Despite conducting these analyses, Clough *et al.* and Earle (2006) provided no rationale to underpin the hypothesised relations between MTQ48 scores and those yielded by the respective subscales measured and did not provide any comprehensive interpretation for their findings. In addition, Earle did not clearly state the specific measures used in his research and Clough *et al.* did not provide sufficient demographic information regarding the sample used in their analyses, which makes it difficult for the reader to draw any robust conclusions regarding these findings. Nonetheless, the patterns of correlations yielded by Earle and Clough *et al.* demonstrate the convergent validity of the MTQ48.

Crust and Swann (2011b) noted the need to examine the convergent validity of the MTQ48 with another measure of MT and achieved this by investigating the relationship between the MTQ48 and the Sport Mental Toughness Questionnaire (SMTQ; Sheard, Golby, & van Wersch, 2009). Findings showed a strong, significant and positive relationship between total MT scores yielded by the MTQ48 and the SMTQ (r = 0.75), and moderate, significant and positive correlations between those conceptually overlapping subscales of the MTQ48 and SMTQ (i.e., MTQ48 confidence – abilities and SMTQ confidence: r = 0.56; MTQ48 control – emotion and SMTQ control: r = 0.49; MTQ48 commitment and SMTQ constancy: r = 0.61). Given that Marsh (2007) suggests that strong correlations (i.e., r > 0.70) are indicative of convergent validity when using instruments purporting to measure the same or substantially overlapping scales, the findings from Crust and Swann's study only provide support for the MTQ48's convergent validity when using Kline's (2005) more liberal guidelines.

The previously discussed research examining the cognitive correlates of the 4/6C's model of MT demonstrates convergent validity for the MTQ48 (see Chapter 2.3.2: Crust & Azadi, 2009, 2010; Crust & Keegan, 2010; Crust & Swann, 2011a; Kaiseler *et al.*, 2009; Nicholls *et al.*, 2008, 2011). When assessing the strength of the correlations relative to Kline's (2005) guidelines (r = 0.30 to 0.50) and the patterns of the correlations obtained in these studies, findings generally demonstrate support for the convergent validity of the MTQ48. However, the previously discussed limitations in the research examining the behavioural correlates of MT (see Chapter 2.3.4: Clough *et al.* 2002; Crust & Clough, 2005) appears to show little support for the MTQ48's criterion validity. Consistent with the construct validation approach (Marsh, 2002), one is required to evaluate the collective evidence which has examined the MTQ48's construct validity. Table 2.5 provides a summary of the evidenced forms of MTQ48 validity and reliability and highlights the growing evidence base to support its convergent validity. Despite some researchers (e.g., Crust & Azadi, 2009; Nicholls *et al.* 2008) positing support for the MTQ48's predictive validity, one may question these conclusions given the cross-sectional nature of their designs (Gignac, 2009). In contrast, the preliminary evidence examining the factorial validity of the MTQ48 coupled with the discrepant internal consistency statistics appears to cast doubt as to the extent at which the MTQ48 items are representing their hypothesised factors. In summary, the evidence pertaining to the MTQ48's construct validity appears to be inconsistent and demonstrates the need for further research to examine the adequacy of the MTQ48 to measure the 4/6C's model of MT.

2.4.2.12 MacKenzie et al.'s Step 10 and the MTQ48

The final aspect of the scale development process requires the researcher to develop norms for the scale to aid the interpretation of score. According to Spector (1992), the meaning of a score can only be determined in relation to a point of reference, which in turn, can be achieved by assessing the distribution of scores in the population of interest. However, at present, norms for the MTQ48 are not available.

2.4.3 *Summary*

Upon evaluating the development of the MTQ48 relative to the scale development literature (e.g., Clark & Watson, 1995; MacKenzie *et al.*, 2011), it is apparent that there are a number of limitations relating to the procedures used by Earle (2006) to develop the MTQ48. Although MacKenzie *et al.'s* framework was not available to Earle when the MTQ48 was developed, the lack of consultation to the available scale development (e.g., Clark & Watson, 1995, Hinkin, 1995; MacKenzie, 2003) and scale validation (e.g., Boomsma, 2000; Marsh, 1997, 2002) literature may explain the poor fit statistics of Gucciardi *et al.'s* (2012) factorial validation.

Table 2.5. Summary of the MTQ48's construct validity.

Type of validity and reliability	Authors	Achievement
Content	Earle (2006)	N
Factorial	Gucciardi *et al.* (2012)	N
	Horsburgh *et al.* (2009)	N
Cronbach's Alpha reliability coefficient	Earle (2006)	Y
	Clough *et al.* (2007)	Y
	Crust (2009)	Y
	Crust & Azadi (2009)	Not presented
	Crust & Azadi (2010)	Not presented
	Crust & Keegan (2010)	N
	Crust & Swann (2011a)	N
	Crust & Swann (2011b)	N
	Kaiseler *et al.* (2009)	N
	Nicholls *et al.* (2008)	N
	Nicholls *et al.* (2011)	N
Test-retest reliability	Earle (2006)	Y
Convergent	Earle (2006)	Y
	Crust (2009)	N
	Crust & Azadi (2009)	Y
	Crust & Azadi (2010)	Y
	Crust & Keegan (2010)	Y
	Crust & Swann (2011a)	Y
	Crust & Swann (2011b)	Y
	Kaiseler *et al.* (2009)	Y
	Nicholls *et al.* (2008)	Y
	Nicholls *et al.* (2011)	Y
Criterion	Clough *et al.* (2002)	N
	Crust & Clough (2005)	N

Note: Y = Yes; N = No

2.5 Summary and Aims of Thesis

The review of literature outlined in this thesis has shown that there is limited support for the factorial validity and reliability of the MTQ48, which may be attributable to poor scale development practices and/or the lack of comprehensive testing of the questionnaire. Given the lack of research which has rigorously tested and reported the factorial validity of the MTQ48 (especially the six factor model, second-order models), one may suggest that researchers may be prematurely dismissing the questionnaire as a valid measure of the 4/6C's model of MT. Consequently, researchers may be prematurely dismissing the 4/6C's model of MT as a valid conceptualisation of MT. Using the scale development literature (e.g., Clark & Watson, 1995; MacKenzie *et al.,* 2011) as a guiding framework, future research is therefore warranted to re-examine the factorial validity of the MTQ48 to provide a clearer picture as to its adequacy in measuring the 4/6C's model of MT. Specifically, research is warranted to add to Gucciardi *et al.'s* (2012) research by re-examining the factorial validity of the hypothesised first-order four factor model of MT. In addition, research is required to extend the literature by examining the factorial validity of the hypothesised six factor model of MT and to examine first- and second-order models of the MTQ48 in an effort to provide a valid tool to measure the 4/6C's model of MT. The provision of a valid tool to measure the 4/6C's model of MT will enable research to examine the traits thought to underpin MT.

Therefore, the objective of this thesis was to test the factorial validity of the MTQ48 to provide a more comprehensive examination than Gucciardi and colleagues (2012) but, on the proviso that this may lead to poor factorial validity, aim to redevelop the measure using more stringent scale development guidelines.

The specific aims of this research were:

1) To examine the factorial validity of the first- and second-order and four and six factor models of the MTQ48 explored in chapter 3.
2) To provide a valid tool to measure the 4/6C's model of MT explored in chapters 4-6.

CHAPTER 3.0

STUDY1: A FACTORIAL ANALYSIS OF THE MENTAL TOUGHNESS QUESTIONNAIRE-48

3.1 Introduction

Over the last decade, a number of instruments have been developed which purport to measure MT in sport (e.g., MTQ48: Earle, 2006; Mental Toughness Inventory: Middleton, 2005; Australian football Mental Toughness Inventory: Gucciardi, Gordon, & Dimmock, 2009; SMTQ: Sheard *et al.*, 2009). To date, the MTQ48 has received the most research utilisation (Gucciardi *et al.,* 2011). In the development of the MTQ48, Earle offered preliminary evidence to support the convergent validity (i.e., presence of significant moderate correlations with other established psychological constructs such as optimism and self-efficacy), internal reliability (i.e., $\alpha > 0.70$ for subscale and total MT scores), and test-retest reliability (i.e., Pearson's correlations ≥ 0.80 over a six week interval). More recent research has provided preliminary support for the convergent validity of the MTQ48 given the presence of significant low-to-moderate correlations with related personality constructs such as optimism and pessimism (e.g., Kaiseler *et al.*, 2009; Nicholls *et al.,* 2008, 2011), and dispositional flow (Crust & Swann, 2011a). Further support for the MTQ48's convergent validity has been offered with other related psychological constructs such as coping effectiveness (e.g., Kaiseler *et al*; Nicholls *et al.,* 2008, 2011) and leadership preferences (Crust & Azadi, 2009).

Despite its abundant research utilisation (e.g., Crust, 2009; Crust & Azadi, 2010), other researchers have expressed reservations regarding the use of the MTQ48 due to the lack of detailed scale development information and factorial validation statistics reported in the literature (e.g., Connaughton & Hanton, 2009; Gucciardi *et al.,* 2011). Close inspection of Clough *et al.'s* (2007) and Earle's (2006) research reveals that although PCA was used in the development and validation of the MTQ48, the CFA measurement model (Marsh, 2007) was not used to test the factorial validity of the *a priori* four and six factor models of the MTQ48 and the second-order representations of these models. Given the importance placed on using CFA to statistically support the hypothesised model (e.g., MacKenzie *et al.,* 2011; Marsh, 2007) and the emphasis placed on establishing factorial validity before any other forms of validity (such as convergent validity) are examined and established (Gignac, 2009; Marsh *et al.*, 2010), one may argue that the preliminary research conducted by Clough and colleagues did not adequately examine the factorial validity of the MTQ48. This in turn, casts doubt regarding the MTQ48's adequacy to measure the 4/6C's model of MT.

Research (e.g., Crust & Swann, 2011a; 2011b) which has adopted the MTQ48 states that the factorial validity of the questionnaire has been independently supported by Horsburgh *et al.*

(2009). Specifically, Horsburgh and colleagues used EFA and CFA to test the factor structure of the MTQ48 and found superior model fit for the first-order four factor solution when compared to the unidimensional single factor solution in a sample of the general population (n = 438). However, inspection of Horsburgh *et al.'s* research reveals that no fit statistics were reported to support their conclusions that "following oblimin rotation, the pattern matrix suggested that the items fit moderately well onto their designated factors" (p. 102) and that "the four factor solution provided a better fit to the data than did a single factor" (p. 102). It is therefore apparent that the information reported by Horsburgh *et al.* does not allow one to inspect the relevant EFA and CFA statistics which in turn, does not enable one to accurately and comprehensively assess the factor structure of the MTQ48. To date, research by Gucciardi *et al.* (2012) is the only study to use CFA to examine the factor structure of the MTQ48. Gucciardi and colleagues examined the first-order four factor model of the MTQ48 via an online survey in two independent samples; a) athletes (n = 686), and b) full-time employees (n = 639). CFA findings revealed that model fit for the first-order four factor solution was unsatisfactory for the athlete (x^2 [1074] = 5511.88, p < 0.001, CFI = 0.487, TLI = 0.462, SRMR = 0.104, RMSEA = 0.078, 90% confidence interval [0.076, 0.080]) and workplace sample (x^2 [1074] = 4928.95, p < 0.001, CFI = 0.521, TLI = 0.497, SRMR = 0.093, RMSEA = 0.075, 90% confidence interval [0.073, 0.077]). Collectively, model fit statistics (fit indices, regression weights) did not support the hypothesised first-order four factor model of the MTQ48. Although preliminary research by Gucciardi *et al.* provides little support for the factorial validity of the MTQ48, a number of limitations were present which may reduce our confidence in their findings. First, Gucciardi *et al.* did not examine the factorial validity of the hypothesised six factor model of the MTQ48. Second, Gucciardi *et al.* did not examine the MTQ48's factorial validity when using second-order models. Finally, Gucciardi *et al.* only examined the factorial validity of the MTQ48 when using an online response method. Given the current confusion surrounding the utility of the MTQ48 as a measure of MT and the previously discussed issues relating to its development (see Chapter 2.4.2), further research is required to address the limitations of Gucciardi *et al.'s* research to shed light on the MTQ48's ability to measure the 4/6C's model of MT.

Construct validation is an ongoing process which requires the adoption of methodologies that demonstrate rigour, reliability, and validity (Marsh, 2007). Re-examinations of an instrument's factor structure are an important consideration in testing the robustness of theoretical models, especially when examining multidimensional constructs across different populations to those adopted in the initial validation process (Gucciardi *et al.*, 2012) as is the case with the MTQ48. Factorial validity has important implications for both practice (e.g., how an instrument is scored, defining subscales based on item content) and theory (e.g., dimensionality, hierarchical representation; Gucciardi *et al.*). Given that recent research (e.g., Gucciardi *et al.*, 2009; Hagger & Chatzisarantis, 2009) has recommended the use of social desirability when assessing inventory validity and that socially desirable responding has been argued to be one of the most prominent sources of systematic error and when present is considered to compromise the validity of participants' responses (Podsakoff *et al.,* 2003), social desirability was measured in Study 1 to assess the degree to which participants were likely to respond to the MTQ48 in a socially desirable manner.

Greater understanding of the factorial validity of the MTQ48 is therefore warranted to provide a comprehensive examination of the MTQ48's adequacy in measuring the 4/6C's model of MT. Consequently, in order to add to and extend the research by Gucciardi *et al.* (2012) and provide a truly comprehensive examination of the MTQ48's factorial validity, the aim of Study 1 was to re-examine the factorial validity of the MTQ48 by using a very large sample of competitive student athletes. Specifically, Study 1 addresses the concerns of Gucciardi *et al.'s* research by a) examining the factorial validity of the hypothesised four (Clough *et al.,* 2002) and six factor (Earle, 2006) models of the MTQ48, b) examining the MTQ48's factorial validity when using first- and second-order models, and c) examining the factorial validity of the MTQ48 when using a paper response method. CFA was deemed to be the most appropriate means of assessing the factor structure of the MTQ48 since it allows researchers to assess the fits of data of an *a priori* model structure and subsequently assess the validity of each item to measure its hypothesised construct (Jöreskog & Sorbom, 1993, 1996; Schutz & Gessaroli, 1993). Given the concerns regarding the adequacy of MTQ48 items to accurately represent the 4/6C's model of MT (see Chapter 2.4.2.5) and the need to fully test the model, model re-specification was conducted where model fit was inadequate.

3.2 Method

Participants

Preliminary data screening showed that the data were nonnormally distributed (see Table 3.1 for univariate normality). The Asymptotic Distribution-free (ADF) estimation method has no distributional assumptions and, theoretically, therefore should perform better in more restrictive conditions of normality (Tabachnick & Fidell, 2001; Byrne, 2010). In an effort to account for the nonnormal distribution in the data, a purposeful recruitment procedure was administered. To successfully perform the ADF method, AMOS stipulates that the sample size has to exceed n*(n+1)/2, where n is the number of observed variables in the model (Arbuckle, 2009). In the present study there were 48 observed variables in the specified model. Consequently, AMOS required a sample in excess of 1,176 participants in order to perform the ADF method. The sample size of this study therefore needed to exceed this value and provided the driving force for the recruitment process. Moreover, this guide ensured that the sample was adequately powered.

Participants (*n* = 1206) were competitive student athletes from six universities in England and Scotland. Twenty-two participants did not fully complete the questionnaire and were subsequently removed from the data analysis. The remaining 1184 participants (*M* age = 20.06 years, *S.D.* = 2.52) consisted of 783 males and 400 females (*M* competitive playing experience in their primary sport = 8.51 years (*S.D.* = 4.09). Five participants did not specify their age and one athlete did not specify their gender. The majority of participants (*n* = 1048) were White British, with the remaining participants (*n* = 133) consisting of a variety of ethnic backgrounds such as White other (*n* = 19), Mixed White and Black Caribbean (*n* = 18), and Black or Black British Caribbean (*n* = 18). Three participants did not specify their ethnicity. Participants were involved in both team sports (*n* = 842), such as football (*n* = 427), rugby (*n* = 129), netball (*n* = 90), and cricket (*n* = 59), and individual sports (*n* = 339), such as athletics (*n* = 87), swimming (*n* = 37), tennis (*n* = 35), and badminton (*n* = 22). Three athletes did not specify their primary sport participation. The highest level of primary sport participation ranged from recreational (*n* = 42), through intervarsity (*n* = 75), club (*n* = 415), county (*n* = 320), regional (*n* = 128), national (*n* = 98) and international (*n* = 73) level. Recreational level was classified as non-structured yet competitive sport participation. Thirty-three participants did not specify their highest level of primary sport participation.

Measures

Mental Toughness Questionnaire (MTQ48)

The MTQ48 (Earle, 2006) is a 48-item inventory which requires respondents to rate their agreement to statements on a 5-point Likert scale ranging from (1) strongly disagree to (5) strongly agree (see Appendix 1.1). The MTQ48 measures six factors of challenge (8 items), commitment (11 items), control – emotion (7 items), control – life (7 items), confidence – abilities (9 items), and confidence – interpersonal (6 items). Example items include “I usually enjoy a challenge” (challenge); “I usually find something to motivate me” (commitment); “I tend to worry about things well before they actually happen” (control - emotion); “I generally feel that I am in control of what happens in my life” (control - life); “I generally feel that I am a worthwhile person” (confidence - abilities); “I usually take charge of a situation when I feel it is appropriate” (confidence - interpersonal). The MTQ48 contains 22 reversed items which are recoded prior to calculation of average scores for each subscale. Items for each factor are totalled and averaged to give a score from 1 to 5 (5 indicating the highest score on the scale). Total mean MT is calculated by totalling the scores from all 48 items and dividing by 48 (see Appendix 1.2 for MTQ48 scoring key).

Social Desirability Questionnaire (SDS)

Due to the demands of recruiting a very large sample, 551 participants were asked to complete the 12-item version of the Marlowe-Crowne Social Desirability scale (Reynolds, 1982; see

Appendix 1.3). Participants were required to rate whether 12 statements regarding personal attributes and traits were true or false to them personally. Items included “There have been times when I was quite jealous of the good fortune of others” and “I have never deliberately said something that hurt someone’s feelings”. Negatively worded items were reversed scored accordingly and responses were summed to give a total social desirability score (ranging from 0-12; see Appendix 1.4 for SDS scoring key). Higher scores are indicative of more social desirability, that is, an individual’s tendency to provide socially acceptable responses.

Procedure

Prior to data collection in all studies of this thesis, institutional ethical approval was obtained (see Appendix 1.5 for ethical approval). Participants were provided with a cover story and informed that the study was investigating the psychological characteristics and thought processes of competitive athletes which required them to complete two questionnaires. They were told that the questionnaires would assess their general psychological attributes. This cover story was implemented in an effort to minimise potential social desirability effects in responding to the MTQ48. Participants who described themselves as athletes currently competing in sport were provided with an information sheet, consent form, athlete demographic questionnaire, the MTQ48, and the SDS (see Appendix 1.6 for participant information sheet, consent form and demographic questionnaire). Participants were recruited over a period of six months. Questionnaires were distributed to athletes during lectures and seminar classes, and were completed in the presence of the author (or a fully briefed assistant) so that any questions could be answered. The questionnaires took approximately 10-15 minutes to complete. Once the questionnaires were completed, participants were thanked for their participation and received a written and verbal debrief explaining the true nature of the study (see Appendix 1.7 for the written debrief).

Data analyses

Data screening

Listwise deletion (n = 22) was employed where missing values were identified (Tabachnick & Fidell, 2001). Data for each observed variable were screened for univariate normality using skewness and kurtosis ratios (Fallowfield, Hale, & Wilkinson, 2005). Analyses identified that the data were skewed and kurtotic (see Table 3.1). Ratios below -2 and above 2 are indicative of univariate nonnormality (Fallowfield *et al.*). Multivariate normality was assessed using the kurtosis critical ratio (C. R.) value. This value represents Mardia’s (1970, 1974) normalised estimate of multivariate kurtosis (Byrne, 2010). Analysis revealed a value of 86.47 which is above Bentler’s (2005) threshold of 5.00 and is therefore indicative of nonnormal data. Data were screened for univariate outliers using box-plots (see Appendix 1.8). Values exceeding three interquartile ranges above the upper quartile range were classified as extreme outliers

(Fallowfield *et al.*). A total of 64 extreme outliers (relating to 47 participants) were identified. In order to assess the impact of the extreme outliers upon the distribution of the overall data set, two independent data sets were formulated. Data set A (unscreened) comprised of the original data set (n = 1,184) and data set B (screened) comprised of the original data set minus the extreme outliers (n = 1,137).

Multivariate outliers were assessed through the computation of the squared Mahalanobis distance (D^2) for each participant. Byrne (2010) argues that D^2 values that sit in isolation and away from all other D^2 values are indicative of multivariate outliers. Examination of Mahalanobis distances identified numerous multivariate outliers (see Appendix 1.9). Although it has been suggested to remove univariate and multivariate outliers when conducting structural equation modelling to ensure the data are normally distributed (Tabachnick & Fidell, 2001), analysis of the distribution of data set A and data set B did not result in any discernible differences in distribution of normality (see Table 3.1 & Appendix 1.10, respectively). Consequently, in light of sample preservation, those cases (n = 47) which were identified as univariate outliers were reinstated. Therefore, all analyses were conducted on the original data set (n = 1184). Descriptive statistics by factor are depicted in Table 3.1. Data screening was conducted using PASW statistics 18.0.

Confirmatory Factor Analysis

CFA tests the robustness of an instrument's factor structure and was used to statistically test the hypothesised models of the MTQ48 (e.g., MacKenzie *et al.*, 2011; Marsh, 2007). CFA is underpinned by a strong theoretical foundation (i.e., *a priori* model) that enables the researcher to evaluate the validity of individual items by determining whether the relationship between each item and its hypothesised latent construct is large and significant (e.g., Jöreskog & Sorbom, 1993, 1996; Schutz & Gessaroli, 1993). Specifically, CFA is achieved by assessing the fit between the reproduced covariance matrix (Σ) and the observed covariance matrix (S). CFA was conducted using AMOS statistics 18.0.

In order to provide a comprehensive examination the MTQ48's factorial validity, first- and second-order models of the four and six factor models of MT were constructed and examined (see Appendix 1.11). The four and six factor models of the MTQ48 had been specified previously by Clough *et al.* (2002) and Earle (2006), respectively. First-order models of MT were constructed to examine the multidimensionality of the proposed four and six factor models (Byrne, 2010). The multidimensionality of the models is represented by the specified inter-correlations between the respective factors (Chen *et al.*, 2005). The second-order models of MT were constructed to examine how the proposed factors might be accounted for by a higher-order MT construct (Chen *et al.*). Although second-order models could be argued to be more stringent and more theoretically driven than first-order models (Byrne, 2010), it was thought that collective analyses would provide a greater pool of evidence from which conclusions can be drawn regarding the MTQ48's factorial validity.

According to Bentler's (1995) six rules for model specification, all identified parameters (known as free model parameters) must be estimated by AMOS (Raykov & Marcoulides, 2000). However, rule six states that for each factor included in a model, its scale metric needs to be calibrated (fixed to a constant) for model estimation to function effectively. Specifically, the metric underlying each factor needs to be defined. First, this was achieved by fixing the path from the second-order MT factor to the first-order challenge factor to 1. Second, one path from each first-order factor to their respective observed variables (items) was fixed to 1. Third, the error terms associated with each observed variable were fixed to 1. Finally, the error terms associated with the first- and second-order factors were fixed to 1 (see Appendix 1.11). Such parameters are called fixed because they do not change value when the model is fit to the observed data (Raykov & Marcoulides, 2000). Due to the nature of the first-order models, the aforementioned first and final step was not implemented.

Estimation method

Once the respective models had been correctly specified in AMOS, CFA was conducted on the first- and second-order four and six factor models of MT. In light of the data being nonnormally distributed, three different estimation methods were utilised; namely the ML, Generalised Least Squares (GLS) and ADF method. This enabled an initial comparison of the fit indices and parameter estimates of the respective methods (see Appendix 1.12). However, the more stringent estimation methods (i.e., GLS, ADF) had difficulty in specifying a proper solution (Blunch, 2008). Despite the sample size exceeding the required threshold, researchers have suggested that unless the sample size is extremely large (1,000 to 5,000 cases: West, Fitch, & Curran, 1995; > 2,500: Tabachnick & Fidell, 2001), the ADF method can perform very poorly and yield severely distorted estimated values and standard errors (Curran, West, & Fitch, 1996; Hu, Bentler, & Kano, 1992; West *et al.*). Consequently, the large sample size used in Study 1 may not have been sufficient to afford the accurate use of the ADF method. Furthermore, the findings from Study 1 support those of Olsson, Foss, Troye, and Howell (2000) in that the ML estimation method provided the best fitting models when compared to GLS and ADF in conditions of nonnormality. The ML estimation method was therefore the primary method used within subsequent model re-specifications. Notwithstanding, in an effort to further examine the factorial validity of the MTQ48, fit indices of the first-order four and six factor models were analysed using the Robust Maximum Likelihood (MLR) estimation method on EQS 6.1 for Windows (Bentler, 2006). The MLR estimation method affords tests of model fit that are robust to nonnormal data (e.g., Beauducel & Herzberg, 2006; Dolan, 1994; Muthén & Kaplan, 1985).

Goodness of fit

Due to the limitations inherent in chi-square likelihood ratio test statistics (e.g., sensitivity to sample size; Byrne, 2010), many researchers (e.g., Byrne, 1998, 2000; Hoyle & Panter, 1995; Kline, 1998; Tanaka, 1993) have suggested using multiple measures of fit indices to provide a more accurate model evaluation process. In Study 1, the following measures were employed: a) CMIN/DF (Wheaton, Muthén, Alwin, & Summers, 1977); b) CFI (Bentler, 1990); c) Parsimonious Comparative Fit Index (PCFI); d) RMSEA (Steiger, 1990); e) Akaike's (1987) Information Criterion (AIC). These fit indices included measures from four different classes (descriptive fit, absolute fit, absolute fit with penalty function, alternative fit). Although some researchers have advocated the use of more conservative thresholds to evaluate model fit (e.g., Hu & Bentler, 1999; Russell, 2002), others have emphasised the need to use more liberal guidelines (e.g., Marsh *et al.*, 2004). Specifically, conservative thresholds indicate good model fit whereas liberal thresholds indicate adequate model fit. The criteria for good fit of the CMIN/DF are values below 2.00 with a non-significant ($p > 0.05$) test result (Byrne, 2010); CFI close to 0.95 (Hu & Bentler, 1999); PCFI close to or above 0.60 (Blunch, 2008); RMSEA below 0.05 indicates good fit, and values below 0.08 indicate adequate fit (Browne & Cudeck, 1993). For completeness, the 90% confidence intervals (90% CI) are provided for the RMSEA. Confidence intervals closely surrounding the RMSEA statistic are indicative of good model fit. For the AIC, lower values are indicative of well-fitting models (Raykov & Marcoulides, 2000). With regards to the fit indices relating to the MLR estimation method, the criteria for good fit of the SRMR are values below 0.05 (Byrne, 2010) and Non-Normed Fit Index (NNFI) above 0.95 (Hu & Bentler. Conversely, the criteria for adequate fit of the CMIN/DF are values below 3.00 with a non-significant ($p > 0.05$) test result (Ullman, 2001); CFI above 0.90 (Bentler, 1992); PCFI close to or above 0.60 (Blunch, 2008); RMSEA below 0.08 (Brown & Cudeck, 1993). Given the presence of conservative and liberal model fit thresholds in the literature, the data gleaned in this thesis will be evaluated using both thresholds.

Regression weights

Regression weights were analysed to indicate the degree to which an item was associated to its respective factor and how much a factor was associated to the second-order MT factor. The current study used standardised regression weights as these are argued to facilitate simpler data interpretation when compared to unstandardised regression weights (Tabachnick & Fidell, 2001). Standardised regression weights usually range from -1 to 1, with values closer to 1 being indicative of stronger relationships (Byrne, 2010). Regression weights which fall outside this range are indicative of an improper factor solution (Blunch, 2008). According to Comrey and Lee (1992), estimates above 0.71 are excellent, 0.63 to 0.70 are very good, 0.55 to 0.62 are good, 0.45 to 0.54 are fair, and 0.32 to 0.44 are considered poor. Regression weights were deemed to be statistically significant at the $p < 0.05$ level.

Model re-specification

The primary aim of Study 1 was to examine the factorial validity of the hypothesised factor structures of the MTQ48. However, in light of the concerns regarding the MTQ48 items to accurately represent the 4/6C's model of MT (see Chapter 2.4.2.5) and the need to fully test the model, model re-specification was employed when there was poor model fit between the sample covariance matrix (*S*) and the estimated covariance matrix (∑). Poor model fit can be due to a number of characteristics including mis-specified correlations between factors, items having low regression weight loadings on their hypothesised factors, and when items inadvertently ask the same question (Byrne, 2010). Modification indices were assessed since they provide the only meaningful information sources regarding CFA model mis-specification (Byrne, 2010). This information is derived from all regression weights and error covariance terms that are fixed to a constant value (in this case 1) in the initial model specification. Specifically, the modification indices relating to the measurement error covariances were analysed. These covariances represent systematic as opposed to random measurement error in item responses which may be attributable to either the items or to the respondents (Aish & Jöreskog, 1990). For instance, errors relating to item characteristics may be indicative of a small omitted factor. Conversely, if the errors represent respondent characteristics, they may reflect response bias (e.g., response acquiescence, social desirability; Aish & Jöreskog, 1990). A high degree of overlap in item content can also result in measurement error covariances. Essentially, this occurs when two items inadvertently ask the same question.

In the event that the hypothesised model structures were not supported by the data, analyses were conducted to generate model re-specifications. The decisions underlying these model re-specifications were directed by identifying high measurement error covariances which sat in isolation and away from all the other modification indices. Such covariances are indicative of mis-specified items (Byrne, 2010). In this study, covariances above 30 were deemed to reflect high measurement error covariances. A progressive item removal protocol was then administered in an effort to re-specify the model. Measurement error covariances between items which were specified to load on the same factor were collated. The content of each item incorporated in the error covariance was compared and assessed to determine whether they were inadvertently asking the same question. In order to address this potential overlap in item content, the item with the lowest relationship (regression weight) with its hypothesised factor was removed from the model (Byrne, 2010). This process was progressive in that once an item had been removed, full CFA analysis was administered so that model fit could be assessed in light of the respective phases of model re-specification. Items were removed in hierarchical order in relation to their error covariances with the highest modification indices being addressed first.

3.3 Results

Descriptive statistics

Descriptive statistics of MTQ48 scores are depicted in Table 3.1. Inter-factor correlations of the MTQ48 were significant and weak to moderate suggesting that the factors represent related yet independent components of the 4/6C's model of MT (see Table 3.2). The correlations between the respective factors and the SDS were generally weak which suggests that socially desirable responding had little impact on the regression weights observed (see Table 3.2). The respective factors of the MTQ48 demonstrated adequate internal reliability ($\alpha > 0.70$: Nunnally & Bernstein, 1994; challenge: 0.70; commitment: 0.77; confidence: 0.78; confidence – abilities: 0.73; confidence – interpersonal: 0.75) except for the control subscales (control: 0.66; control – emotion: 0.53; control – life: 0.61). Overall, mean total MT scores demonstrated high internal reliability ($\alpha = 0.89$; see Table 3.2).

Confirmatory Factor Analyses

Goodness of fit

The CFA's conducted on the hypothesised model specifications revealed poor model fit relative to the conservative and liberal guidelines (see Study 1 Method). Table 3.3 indicates that although the six factor models provided marginally better model fit than the four factor models, the respective fit indices are generally below the conservative thresholds (identified in parentheses) and liberal thresholds. Fit indices of the first-order four (Satorra-Bentler $\chi^2(1074) = 5618.337$, $p < .001$, CFI = 0.617, NNFI = 0.597, SRMR = 0.070, RMSEA = 0.060 [0.052, 0.057]) and six factor models (Satorra-Bentler $\chi^2(1065) = 4865.255$, $p < .001$, CFI = 0.679, NNFI = 0.661, SRMR = 0.068, RMSEA = 0.055, [0.055, 0.061]) revealed no discernible differences to those models examined using the ML estimation method in AMOS.

Regression weights

Four factor models

The four factor second-order model revealed that the control factor was found to have the strongest relationship with the second-order MT factor and commitment was found to have the weakest relationship with the second-order MT factor. However, in line with the guidelines proposed by Comrey and Lee (1992), all four factors could be considered excellent in terms of their relationship with the second-order MT construct. Regression weights of the first- and second-order four factor models revealed that in general, items relating to the commitment factor were found to have the strongest relationships with their hypothesised factor and items relating to the control factor were found to have the weakest relationships with their hypothesised factor (see Appendix 1.13).

Inspection of item regression weights shows that only 24 out of 48 items (50%) could be considered to have fair or above relationships with their hypothesised factors. Specifically, estimates obtained from individual items of the challenge and commitment factors were generally considered to have fair to good relationships with their hypothesised factors, with 5 out of 8 items (62.5%) and 8 out of 11 items (72.7%) considered fair or above, respectively. However, no items from the challenge and commitment factors were considered to have very good or above relationships with their hypothesised factors. Estimates obtained from the individual items of the control and confidence factors were considered to have poor relationships with their hypothesised factors, with 9 out of 14 items (64.2%) and 9 out of 15 items (60%) considered poor or below, respectively. All items showed a statistically significant relationship ($p < 0.05$) with their hypothesised factor except item MTQ26R (see Appendix 1.13).

Six factor models

Regression weights of the six factor second-order model revealed that the control - life factor was found to have the strongest relationship with the second-order MT factor and the confidence - interpersonal factor was found to have the weakest relationship with the second-order MT factor. Inspection of factor estimates revealed that 5 out of the 6 factors in the six factor second-order model could be considered excellent in terms of their relationship with the second-order MT construct, whereas confidence-interpersonal could be considered good. Analysis of the first- and second-order six factor models revealed that in general, the items relating to the commitment and the confidence – interpersonal factors were found to have the strongest relationships with their respective factors and that the items relating to the control factor were found to have the weakest relationships with their respective factors (see Appendix 1.13).

Only 32 out of 48 items (66.6%) could be considered to have fair or above relations to their hypothesised factors. Specifically, estimates obtained from individual items of the challenge, commitment, confidence – abilities, and confidence - interpersonal factors were generally considered to have fair to good relationships with their hypothesised factors, with 5 out of 8 items (62.5%), 8 out of 11 items (72.7%), 7 out of 9 items (77.7%) and 6 out of 6 items (100%) considered fair or above, respectively. However, only two items from the confidence - interpersonal factor (MTQ43, MTQ38) were considered to have very good or above relationships with their hypothesised factors. Estimates obtained from the individual items of the control - emotion and control - life factors were considered to have poor relationships with their hypothesised factors, with 4 out of 7 items (57.1%) considered poor or below for both factors. Only item MTQ26R and item MTQ34 failed to display statistically significant relationships ($p < 0.05$) with their hypothesised factors (see Appendix 1.13).

Model re-specification

Due to the poor fit of the respective hypothesised models of MT and the need to fully examine the 4/6C's model of MT, modification indices were analysed to guide model re-specification in an effort to improve model fit. Independent model re-specification protocols and their respective fit indices are presented in Appendix 1.14. Table 3.4 summarises the fit indices of the best fitting revised models. Findings show that although the four factor first-order revised model showed the highest level of change in CFI (ΔCFI) from its respective hypothesised model specification, the six factor first-order revised model provided the best fitting model. However, the improvements in model fit failed to provide an adequately fitting model in accordance with the conservative and liberal thresholds.

Table 3.1. Descriptive statistics of the MTQ48 by factor.

Factor		Item	Mean	*S.D.*	S ratio	K ratio
Challenge	MTQ4	Challenges usually bring out the best in me	4.10	0.79	-8.11	-0.26
	MTQ6R	Unexpected changes to my schedule generally throw me	3.26	0.99	-5.23	-2.88
	MTQ14R	I often wish my life was more predictable	3.56	1.04	-6.58	-2.79
	MTQ23	I generally cope well with any problems that occur	3.66	0.68	-8.16	2.79
	MTQ30	I am generally able to react quickly when something unexpected happens	3.63	0.73	-9.42	5.68
	MTQ40	I usually look forward to changes in my routine	3.27	1.02	-3.82	-3.20
	MTQ44	I usually enjoy a challenge	4.24	0.71	-10.81	5.71
	MTQ48	I can usually adapt myself to challenges that come my way	3.95	0.65	-8.11	9.32
Commitment	MTQ1	I usually find something to motivate me	4.04	0.72	-8.98	6.69
	MTQ7	I don’t usually give up under pressure	4.09	0.84	-14.88	9.72
	MTQ11R	“I just don’t know where to begin” is a feeling I usually have when presented with several things to do at once	3.05	1.15	-1.69	-6.64
	MTQ19	I can generally be relied upon to complete the tasks I am given	4.11	0.68	-9.28	8.75
	MTQ22R	I am easily distracted from tasks that I am involved with	3.11	1.06	-3.54	-5.13
	MTQ25	I generally try to give 100%	4.35	0.74	-15.39	9.07
	MTQ29R	When faced with difficulties I usually give up	4.19	0.82	-17.19	14.49
	MTQ35R	I usually find it difficult to make a mental effort when I am tired	2.58	1.04	5.64	-4.57
	MTQ39	I can normally sustain high levels of mental effort for long periods	3.23	0.96	-2.96	-4.04
	MTQ42R	I usually find it hard to summon enthusiasm for the tasks I have to do	3.68	0.91	-8.22	-0.35
	MTQ47R	When I face setbacks I am often unable to persist with my goal	3.59	0.95	-7.68	-1.76
Control - emotion	MTQ21R	I generally find it hard to relax	3.62	1.11	-8.33	-2.85
	MTQ26R	When I am upset or annoyed I usually let others know	2.98	1.19	0.56	-6.82
	MTQ27R	I tend to worry about things well before they actually happen	2.72	1.18	3.79	-6.39
	MTQ31	Even when under considerable pressure I usually remain calm	3.48	0.88	-5.29	-1.86
	MTQ34	I generally hide my emotion from others	3.17	1.14	-1.89	-6.17
	MTQ37R	When I am feeling tired I find it difficult to get going	2.57	0.99	6.79	-3.39
	MTQ45	I can usually control my nervousness	3.40	1.03	-7.22	-2.28

Table 3.1. *(continued).*

Factor		Item	Mean	*S.D.*	S ratio	K ratio
Control - life	MTQ2	I generally feel in control	3.92	0.66	-10.26	11.22
	MTQ5	When working with other people I am usually quite influential	3.65	0.77	-2.22	-1.97
	MTQ9R	I usually find myself just going through the motions	3.17	0.96	-1.53	-3.14
	MTQ12	I generally feel that I am in control of what happens in my life	3.84	0.79	-10.04	5.52
	MTQ15R	Whenever I try to plan something, unforeseen factors usually seem to wreck it	3.41	0.95	-6.40	-1.56
	MTQ33R	Things just usually happen to me	3.33	0.96	-2.54	-2.35
	MTQ41R	I feel that what I do tends to make no difference	3.81	0.83	-8.14	2.07
Confidence - ability	MTQ3	I generally feel that I am a worthwhile person	4.09	0.73	-9.70	7.56
	MTQ8	I am generally confident in my own abilities	3.87	0.80	-10.50	5.49
	MTQ10R	At times I expect things to go wrong	2.78	1.01	3.82	-3.90
	MTQ13	However bad things are, I usually feel they will work out positively in the end	3.74	0.91	-7.11	-0.99
	MTQ16	I generally look on the bright side of life	4.01	0.80	-12.34	8.62
	MTQ18R	At times I feel completely useless	3.61	1.03	-6.67	-3.12
	MTQ24	I do not usually criticise myself even when things go wrong	2.34	0.98	7.85	-1.38
	MTQ32R	If something can go wrong, it usually will	3.55	0.95	-6.37	-1.50
	MTQ36R	When I make mistakes I usually let it worry me for days after	3.13	1.15	-1.59	-6.84
Confidence - interpersonal	MTQ17	I usually speak my mind when I have something to say	3.73	0.94	-6.82	-2.84
	MTQ20	I usually take charge of a situation when I feel it is appropriate	3.94	0.80	-8.60	2.94
	MTQ28R	I often feel intimidated in social gatherings	3.79	1.06	-9.89	-0.95
	MTQ38	I am comfortable telling people what to do	3.55	0.95	-7.42	-0.83
	MTQ43	If I feel somebody is wrong, I am not afraid to argue with them	3.55	1.04	-7.99	-1.72
	MTQ46R	In discussions, I tend to back-down even when I feel strongly about something	3.79	1.02	-9.04	-1.45

Note: S ratio = Skewness ratio; K ratio = Kurtosis ratio.

Table 3.2. Correlations between factors of the MTQ48 and SDS scores and Cronbach's Alpha reliability coefficients.

	MT factors								
	CHAL	COM	CONT	CE	CL	CONFID	CA	CI	MTMT
CHAL	(0.70)								
COM	0.49**	(0.77)							
CONT	0.54**	0.54**	(0.66)						
CE	0.41**	0.36**	0.84**	(0.53)					
CL	0.47**	0.53**	0.76**	0.32**	(0.61)				
CONFID	0.56**	0.48**	0.61**	0.39**	0.61**	(0.78)			
CA	0.51**	0.45**	0.62**	0.46**	0.56**	0.85**	(0.73)		
CI	0.40**	0.32**	0.34**	0.14**	0.43**	0.76**	0.32**	(0.75)	
MTMT	0.74**	0.76**	0.84**	0.61**	0.75**	0.85**	0.77**	0.58**	(0.89)
SDS	0.20**	0.33**	0.22**	0.19**	0.17**	0.04	0.18**	-0.13**	0.22**

Note: **Significance is at $p < 0.01$. MTQ48 inter-factor correlations and internal reliability coefficients are based on $n = 1184$, whereas correlations between the MTQ48 and SDS are based on $n = 551$. Cronbach's Alpha reliability coefficients are displayed in parentheses. CHAL = Challenge; COM = Commitment; CONT = Control; CE = Control - emotion; CL = Control - life; CONFID = Confidence; CA = Confidence - ability; CI = Confidence - interpersonal; MTMT = Mean Total MT; SDS = Social Desirability Scale.

Table 3.3. Summary of fit indices across hypothesised model specification.

	Fit indices					
CFA model Criterion values	CMIN/DF (< 2.00)	*P* (> 0.05)	CFI (> 0.95)	PCFI (> 0.60)	RMSEA (90% CI) (< 0.05)	AIC (lower = better)
First-order four factor	6.129	0.000	0.614	0.585	0.066 (0.064, 0.067)	6786.535
First-order six factor	5.334	0.000	0.677	0.639*	0.061 (0.059, 0.062)	5902.412
Second-order four factor	6.165	0.000	0.611	0.583	0.066 (0.065, 0.068)	6929.326
Second-order six factor	5.433	0.000	0.667	0.635*	0.061 (0.060, 0.063)	6134.950

Note: * denotes good fit.

Table 3.4. Summary of fit indices across best fitting revised model specification.

	Fit indices						
CFA model Criterion values	CMIN/DF (< 2.00)	*P* (> 0.05)	CFI (> 0.95)	Δ CFI	PCFI (> 0.60)	RMSEA (90% CI) (< 0.05)	AIC (lower = better)
First-order four factor	5.480	0.000	0.788	0.174	0.723*	0.062 (0.059, 0.064)	2318.705
First-order six factor	4.723	0.000	0.798	0.121	0.731*	0.056 (0.054, 0.058)	2744.062
Second-order four factor	5.540	0.000	0.760	0.149	0.707*	0.062 (0.060, 0.064)	2926.319
Second-order six factor	5.155	0.000	0.783	0.116	0.727*	0.059 (0.057, 0.061)	2901.665

Note: * denotes good fit.

3.4 Discussion

The aim of Study 1 was to examine the factorial validity of the MTQ48 using a sufficiently large student athlete sample. The findings of the respective CFA's provided little support for the hypothesised models of the MTQ48 in that fit indices of the first- and second-order, four and six factor models revealed inadequate model fit when using both conservative and liberal guidelines (see Table 3.3). The findings of Study 1 are consistent with those obtained by Gucciardi *et al.* (2012) whereby little support was provided for the first-order four factor model of the MTQ48. The inadequate fit indices obtained in Study 1 adds to the findings of Gucciardi *et al.* which suggest that in its current form, the MTQ48 is not a valid measure of the 4/6C's model of MT which it intends to capture. Inspection of regression weights provided further evidence to question the factorial validity of the MTQ48 in that associations between the items and their respective factors were relatively weak. For the four and six factor models (first- and second-order), only 24 out of 48 items (50%) and only 32 out of 48 items (66.6%) could be considered fair or above, respectively (Comrey & Lee, 1992). Results also indicated that there were very few items which could be considered to have very good to excellent relationships with their hypothesised factors. For the four factor models, no items reached this threshold. Similarly, for the six factor models only two items from the confidence-interpersonal factor (MTQ38, MTQ43) reached this threshold.

The findings from Study 1 are comparable to those observed by Gucciardi *et al.* (2012) in that only 22 out of 48 items in the first-order four factor model could be considered as fair or above and only 5 out of 48 items could be considered very good or above, respectively (Comrey & Lee, 1992). Further inspection of regression weights in Study 1 identified five items which had particularly low relationships with their hypothesised factors across all models examined. These items included three items from the control - emotion factor (MTQ26R, MTQ34, MTQ37R), one item from the control - life factor (MTQ9R), and one item from the confidence - abilities factor (MTQ24R). According to Comrey and Lee (1992), these values could be considered too low to be interpreted, which may provide evidence to suggest that these items are major contributors to the lack of support found for the hypothesised models examined in Study 1. Similarly, Gucciardi *et al.* identified the aforementioned items as having extremely low relationships with their hypothesised factors and identified a further 12 items which could be considered too low for interpretation. Specifically, Gucciardi *et al.* identified items relating to all four factors examined; challenge (MTQ6R, MTQ14R), commitment (MTQ35R), control (MTQ15R, MTQ33R, MTQ21R, MTQ27R) and confidence (MTQ10R, MTQ28R, MTQ32R, MTQ36R, MTQ46R). The findings of Study 1 coupled with those found by Gucciardi *et al.* therefore indicate that there are a number of items relating to all four MT factors which are inadequate representations of

their hypothesised factors, with the most pronounced inadequate items relating to the control and confidence factors.
In addition, item MTQ34 demonstrated a negative association with its control - emotion factor. This finding is similar to that observed by Crust and Swann (2011b) and Gucciardi *et al.* (2012) in that item MTQ34 was found to be unrelated to control - emotion. One possible reason for this unexpected negative association might be due to the content of item MTQ34. This item reflects the facet of the control - emotion factor which argues that mentally tough performers are "less likely to reveal their emotional state to other people" (Clough *et al.,* 2007, p. 4; see Table 2.1). Although this facet of control - emotion may appear intuitively compelling, Clough *et al.* fail to provide an empirical rationale for its inclusion. Nicholls and Polman's (2007) systematic review of coping in sport would appear to partially conflict with this supposition in that findings generally appear to highlight the importance of being able to effectively control one's emotions to maintain/enhance performance as opposed to supporting the use of covert coping strategies to suppress one's emotions. Based upon Nicholls and Polman's review, an athlete's MT levels may be more accurately determined by their effectiveness in controlling emotions as opposed to whether one controls their emotions covertly or overtly. The ambiguity surrounding item MTQ34 and its apparent negative association with its hypothesised control - emotion factor suggests that this item warrants major revision.

Although Study 1 provided little support for the respective models of the MTQ48, inspection of the fit indices and regression weights revealed that the six factor models provided a marginally better fitting solution when compared to the four factor models (see Table 3.3 for fit indices & Appendix 1.13 for regression weights). This finding suggests that the MTQ48 is better understood when underpinned by Earle's (2006) 6C's model of MT whereby control and confidence are subdivided into two nested components as opposed to Clough *et al.'s* (2002) 4C's model of MT. Inspection of the tested models also revealed that the first-order models provided marginally superior fit statistics when compared to their respective second-order models (i.e., four factor first-order model vs. four factor second-order model). This finding suggests that the MTQ48 is better understood when its underpinning latent constructs are hypothesised to be correlated with the remaining constructs in the model. The marginal superiority of the first-order models in Study 1 are therefore inconsistent with the theoretical predictions of hardiness (Kobasa, 1979) and previous hardiness scales (e.g., Hystad, Eid, Johnsen, Laberg, & Bartone, 2010) which argue that hardiness is best understood when its underpinning constructs are accounted for by a higher-order hardiness construct. This finding is also inconsistent with other MT instruments (e.g., Gucciardi & Gordon, 2009; Sheard *et al.*, 2009) which have supported the notion of a higher-order representation of MT.

In an effort to extend the research of Gucciardi *et al.* (2012), Study 1 used model re-specification protocols to further examine the factorial validity of the MTQ48. Examination of the respective independent model re-specifications revealed that model fit did improve across all models analysed, yet did not collectively reach the required thresholds (both conservative and liberal) to adequately support the fit of the respective revised models of the MTQ48 (see Table 3.4 & Appendix 1.14). The findings appear to suggest that despite rigorous model re-specification protocols, the revised models of the MTQ48 remain unacceptable. Despite a growing body of evidence supporting the MTQ48's convergent validity, the findings of Study 1 emphasise the importance of assessing an instrument's factorial validity prior to assessing other forms of validity (Gignac, 2009; Marsh *et al.,* 2010). Indeed, the utility of an instrument is underpinned by the degree at which its items accurately capture the constructs it intends to capture (e.g., MacKenzie *et al.* 2012; McGrath, 2005).

A possible explanation for the poor fit of the models examined could have been due to the extent to which the items represent their hypothesised factor. As reflected by the poor regression weights obtained, it appears that a large proportion of items are inadequately representing their hypothesised factor definitions forwarded by Clough *et al.* (2007). As previously discussed in Chapter 2.4.2.5, inspection of MTQ48 item content (face validity) revealed 18 out of 48 items to be poor representations of their hypothesised factor definitions (see Table 2.4). Inspection of item content provides credence to this argument. For instance, item MTQ9R – "I usually find myself just going through the motions" does not appear to accurately represent the hypothesised definition of the control - life factor which states that individuals high in control – life have the ability to control one's life, feel that their plans will not be thwarted and that they can make a difference (see Table 2.1 for control – life definition). Similarly, item MTQ37R – "When I am feeling tired I find it difficult to get going" does not appear to accurately reflect the control – emotion factor which encapsulates an individual's ability to control their emotions, keep their anxieties in check and be less likely to reveal their emotional state to other people (see Table 2.1 for the control – emotion definition). One may therefore argue that the MTQ48 is not an accurate representation of the 4/6C's model of MT which it purports to measure. Consequently, it appears that Earle (2006) overlooked a critical step in the scale development process whereby the ultimate objective of item generation is to develop an item pool that encapsulates the core facets of the focal construct (MacKenzie *et al.,* 2011).

A further explanation for the poor fit indices obtained could have been due to the structure and clarity of the items. The scale development literature (e.g., Clark & Watson, 1995; MacKenzie *et al.,* 2012) provides guidelines to inform item wording and structure that emphasise the importance of ensuring clarity, specificity, and brevity with each item. Specifically, Clark and Watson suggest that the exact phrasing of items can exert a profound influence on the construct being measured. However, inspection of the MTQ48 indicates that a considerable amount of items do not adequately fulfil these criteria. For instance, item MTQ11R – "'I just don't know where to begin' is a feeling I usually have when presented with several things to do at once" does not appear to share the same structure (i.e., the use of a quotation) as the remaining MTQ48 items (see Appendix 1.1). Similarly, item MTQ16 - "I generally look on the bright side of life" and item MTQ33R – "Things just usually happen to me" appear to be quite vague and thus may have confused the respondent in terms of what the question is really asking. In addition, item MTQ26R – "When I am upset or annoyed I usually let others know" contains double barrelled wording which conflicts with the recommendations of the scale development literature (e.g., Clark & Watson, 1995; MacKenzie *et al.* 2011). One may therefore argue that the lack of clarity and specificity surrounding the items of the MTQ48 may have made it difficult for the respondents to accurately answer the question which in turn, may have contributed towards the poor factorial validity of the models examined.

Another explanation for the poor fit of the models examined could have been due to the non-normal distribution of data in that analyses revealed numerous univariate and multivariate outliers (see Appendix 1.8 & Appendix 1.9, respectively). In order to comprehensively assess the hypothesised models of the MTQ48 in light of the data being nonnormal, the ML, GLS, and ADF estimation methods were utilised. Although the sample size satisfied AMOS's criteria to run the ADF method, it may not have been sufficient to yield accurate estimates (Curran *et al.*, 1996; Hu *et al.*, 1992; West *et al.*, 1995). Given that the ML method does not explicitly account for nonnormal data, the first-order four and six factor models of the MTQ48 were re-analysed using the MLR estimation method on EQS 6.1 for Windows (Bentler, 2006). However, findings revealed no discernible differences in model fit statistics across the ML and MLR estimation methods, which in turn, provides support for the adequacy of the ML estimation method in conditions of nonnormality (e.g., Olsson *et al.*, 2000).

The use of a cover story and social desirability scale could be considered a strength of this study. Recent research (e.g., Gucciardi *et al.*, 2009; Hagger & Chatzisarantis, 2009) has called for the use of social desirability scales to assess the propensity of a measure to afford socially desirable responding. The weak correlations observed in this study suggest that minimal socially desirable responding of the MTQ48 was present in this sample. Future scale development research should consider the impact of social desirability when constructing and validating a scale and should not overlook the use of a cover a story to minimise the adverse effects of social desirability. In addition, the large sample size utilised in Study 1 encompassed a range of student athlete demographic attributes (e.g., sports played, number of years competitive playing experience). Although the sample was limited to student athletes, the range of athletic abilities captured was indeed diverse and not limited to collegiate playing levels given that the distribution of athletes' highest level of competitive experience was relatively balanced, ranging from international to recreational (see Participants section for demographic information). However, one limitation of Study 1 could be that the sample used was heavily weighted by White British competitive student athletes. Despite participants being recruited from six Universities across England and Scotland, the ethnic backgrounds captured was indeed quite narrow (see Participants section).

The findings of Study 1 coupled with those observed by Gucciardi *et al.* (2012) suggest that the factorial validity of the MTQ48 is poor in its current format. Study 1 addresses Gucciardi *et al.'s* call for future research to examine the factorial validity of the MTQ48 in a larger sample of athletes in that the sample size used in this study could be considered excellent (> 1,000: Comrey & Lee, 1992). Findings suggest that significant revision of the instrument is needed in order to improve its factor structure in that systematic model re-specification protocols were not sufficient to combat the inadequacies of the originally hypothesised models of the MTQ48. Although a number of explanations have been offered to explain the poor factorial validity of the MTQ48, the data suggests that there are major concerns regarding the adequacy (face validity) of the MTQ48 items to represent Clough *et al.'s* (2007) factor definitions of the 4/6C's model of MT. Previous studies using the MTQ48 (e.g., Crust & Azadi, 2009; Kaiseler *et al.*, 2009; Nicholls *et al.*, 2008) should therefore be interpreted with caution given the lack of support for the factorial validity of the MTQ48 using both online (Gucciardi *et al.*, 2012) and paper formats (Study 1) of the instrument.

3.5 Summary

The aim of Study 1 was to examine the factorial validity of the MTQ48 using a sufficiently large student athlete sample. The findings revealed little support for the factor structure of the hypothesised models and the revised models of the MTQ48. The findings therefore indicate that the MTQ48 is an inadequate measure of the 4/6C's model of MT. Collective poor fit indices, weak regression weights, and poor item face validity suggest that research is required to develop items which better represent the factor definitions forwarded by Clough *et al.* (2007) to underpin the 4/6C's model of MT. Study 2 will aim to achieve this by extending the findings of Study 1 by developing new items to better represent the 4/6C's model of MT in an effort to provide a valid instrument to assess the 4/6C's model of MT.

CHAPTER 4.0

STUDY 2: THE DEVELOPMENT AND FACTORIAL ANALYSIS OF THE UNIVERSITY OF CHICHESTER MENTAL TOUGHNESS QUESTIONNAIRE

4.1 Introduction

Many researchers have expressed continual concerns regarding the factorial validity of the MTQ48 (e.g., Connaughton & Hanton, 2009; Gucciardi *et al.*, 2011). Chapter 2 discussed a number of concerns with the scale development procedures used to construct the MTQ48 and highlighted problems with the content and structure of its items. Specifically, inspection of MTQ48 item face validity revealed problems with their content in that 18 out of the 48 items were deemed to be poor representations of their respective factor definitions (see Chapter 2.4.2.5 & Table 2.4). In addition, inspection of MTQ48 item face validity revealed a number of problems with their structure relative to the guidelines outlined in the scale development literature (e.g., Clark & Watson, 1995; MacKenzie *et al.*, 2011; see Chapter 2.4.2.5). Study 1 provided support for concerns regarding the MTQ48's ability to measure the 4/6C's model of MT. Given the inadequate fit indices and weak regression weights obtained in Study 1 and by Gucciardi *et al.* (2012), it was suggested that poor item representation appears to provide a strong explanation for the lack of factorial validity of the MTQ48. According to Step 2 of MacKenzie *et al.'s* scale development framework, the ultimate objective of the item generation process is to develop an item pool that encapsulates the core facets of the focal construct. The evidence yielded in Study 1 and by Gucciardi *et al.* therefore questions the item generation procedures adopted by Earle (2006) to construct the MTQ48.

It is important to note, however, that the evidence yielded and the associated interpretations of this evidence should only be used to assess the validity of the MTQ48 as opposed to a means of evaluating the utility of the 4/6C's model of MT. Research which uses the information gleaned from Study 1 and by Gucciardi *et al.* (2012) to conclude that the 4/6C's model of MT has limited utility may prematurely dismiss the model as a valid conceptualisation of MT. In order to fully test the utility of the 4/6C's model of MT, research is required to redevelop the MTQ48 by regenerating new items which better represent the hypothesised factor definitions forwarded by Clough *et al.* (2007) to underpin the 4/6C's model of MT. Using the scale development literature as a guiding framework (e.g., Clark & Watson, 1995; MacKenzie *et al.*, 2011), the first aim of Study 2, therefore, was to regenerate items which better represent the 4/6C's model of MT in an effort to develop a valid measure of MT, provisionally named the University of Chichester Mental Toughness Questionnaire (UCMTQ). The second aim of Study 2 was to examine the factorial validity of the UCMTQ. It is hypothesised that the first- and second-order four and six factor models of the UCMTQ will provide adequately fitting models. Based on the findings of

Study 1, it is hypothesised that the first-order models and the six factor models will provide better fitting models than the second-order models and four factor models, respectively.

4.2 Method

Participants

When conducting factor analysis, Comrey and Lee (1992) classify sample sizes above 300 as good and sample sizes above 500 as very good. Moreover, Stevens (1996) suggests using five participants per variable to obtain adequate power. Given that the UCMTQ consisted of 45 items, a minimum sample of 225 participants was needed. Using these classifications as a guide, the target sample for Study 2 was set at 500. Participants were 505 competitive student athletes from two universities in Southern England. Twenty-five did not fully complete the questionnaire and were subsequently removed from the data analysis. The remaining 481 participants (*M* age = 20.20 years, *S.D.* = 2.15) consisted of 321 males and 160 females with a mean of 11.48 years (*S.D.* = 4.09) competitive playing experience in their primary sport. The vast majority of participants (93.6%) were White British, with the remaining participants (*n* = 30) consisting of a variety of ethnic backgrounds such as Mixed other (*n* = 6), Black or Black British African (*n* = 5), Mixed White and Black Caribbean (*n* = 4), and White Irish (*n* = 4). Two participants did not specify their age and one participant did not specify their ethnicity. Participants were involved in both team sports (*n* = 365), such as football (*n* = 193), rugby (*n* = 53), netball (*n* = 30), and cricket (*n* = 30), and individual sports (*n* = 116), such as athletics (*n* = 21), combat sports (*n* = 14), and tennis (*n* = 14). The highest level of primary sport participation ranged from recreational (*n* = 7) through intervarsity (*n* = 38), club (*n* = 173), county (*n* = 150), regional (*n* = 46), national (*n* = 46), and international (*n* = 21) level.

Stem generation and scaling

The stem generated for the UCMTQ was an extension of the stem used in the MTQ48. Previous questionnaires measuring traits (e.g., Personal Views Survey III-R: Maddi & Khoshaba, 2001) have incorporated an overarching preceding statement to each item to clarify how the respondents are required to complete the questionnaire. Accordingly, “In general” was added to the stem of the UCMTQ. In an effort to further enhance the clarity of the stem, an example item was incorporated to illustrate how participants should read each item (see Appendix 2.1).

Scaling is a dubious subject in the scale development literature with populist articles (e.g., Clark & Watson, 1995; Comrey, 1988) advocating the use of a variety of response formats (Likert, visual analogue, dichotomous). The ambiguity surrounding scaling is reflected by Clark and Watson who stated that “we cannot conclude that one type of format is generally preferable to the other” (p. 313). From the recommendations, however, it is strongly recommended that a proposed format should be pilot tested to ascertain preliminary information about both response reactions and response option distributions. Accordingly, it was decided to pilot the response format proposed by Earle (2006) for the MTQ48. The MTQ48 uses a 5-point Likert scale anchored by strongly disagree (1) to strongly agree (5). Postgraduate students (*n* = 2) who exhibited knowledge of the MT and scale development literature were used to evaluate the readability, clarity, and adequacy of the proposed response format. Participants deemed the response format adequate for use and no structural changes were made.

Item generation

According to Mackenzie *et al.* (2011), the ultimate goal of the item generation process is to produce a pool of items that fully captures the varying facets of the target construct while minimising the extent to which items tap into concepts other than the one under examination. In line with Step 3 of Mackenzie *et al.'s* scale development framework, a series of steps were used to guide the initial item generation process. First, items were extracted from a variety of sources including reviews of the MT literature, deduction from the theoretical definitions of the 4/6C's model of MT, suggestions from experts in the field of MT and scale development, and an examination of existing measures relating to the target construct(s). A range of questionnaires were used from a variety of related domains including MT (e.g., MTQ48: Earle, 2006; SMTQ: Sheard *et al.*, 2009), hardiness (e.g., Personal Views Survey III-R: Maddi, 1997; Revised Dispositional Resilience Scale: Bartone, 1995), locus of control (e.g., Rotter's Locus of Control Scale; Rotter, 1966), personality (NEO-Five Factor Inventory; McCrae & Costa, 2004), stress (e.g., Perceived Stress Scale: Cohen, Karmarack, & Mermelstein, 1983; Stress Appraisal Measure: Peacock & Wong, 1990), self-confidence (e.g., General Self-efficacy Scale: Schwarzer & Jerusalem, 1995), self-esteem (e.g., Rosenberg Self-esteem Scale: Rosenberg, 1989), anxiety (Sport Anxiety Scale: Smith, Smoll, & Schutz, 1990), and resilience (e.g., Resilience Scale: Wagnild & Young, 1993; Connar-Davidson Resilience Scale: Connor & Davidson, 2003).

In order to successfully capture all the facets of each target factor definition, the first author analysed the extent to which the collated items represented the respective factor definitions. In the instance where the factor definition was not sufficiently represented, the author adapted existing items or generated new items. At this stage, it has been suggested that the wording of all items should be considered and analysed (Peterson, 2000; Podsakoff *et al.*, 2003; Spector, 1992; Torangeau *et al.*, 2000) to ensure that item wording is as simple and precise as possible. An iterative process of item scrutiny was conducted by the author and supervisors (primary and secondary) to assess the wording and clarity of items. To systematically assess item clarity, a number of areas have been suggested to be considered (MacKenzie *et al.*, 2011). First, doubled barrelled items (e.g., I am confident in my skills and abilities) were split into two single-idea statements, or if that proved to be problematic in terms of item overlap, one facet of the initial item was retained and the second facet of the item was removed. Second, ambiguous or unfamiliar language was simplified and made more specific and concise. Finally, efforts were made to refine or remove any items which were deemed to facilitate obvious social desirability responding.

Cooper (2010) suggests that there should be consistency in the length, tone, and structure at which items are phrased to further enhance the coherency of items. As a result, the first author analysed these facets and refined nonconforming items accordingly. Once the items had been systematically revised, the first author and supervisors employed a further iterative refinement procedure to identify the top 10 most representative items for each respective factor of the 4/6C's model of MT. This was achieved by deductively examining each item with respect to its hypothesised factor definition. Through this process, item orientation was considered. Cooper suggests that an item pool should exhibit an equal split of positively worded and negatively worded items. Negatively worded items have been suggested to be an imperative inclusion in the item generation process in that they provide a remedial solution to potential response acquiescence (i.e., yea-saying or nay-saying: Clark & Watson, 1995; Cooper; MacKenzie *et al.*, 2011). An initial pool of 60 items was generated, with each factor being represented by 10 items.

Item validation

Using Step 3 of MacKenzie *et al.'s* (2011) scale development model as a guiding framework, an item validation task was constructed to assess item content. The item validation task was piloted using the same postgraduates (*n* = 2) who piloted the stem and required participants to evaluate the items for readability, clarity and adequacy. Minor structural changes were made. Subsequent to this, in line with previous scale development research in sport and exercise psychology (e.g., Bartholomew *et al.*, 2010; Williams & Cumming, 2011), an independent group (*n* = 6) of expert scrutineers were recruited to review the adequacy of the initial 60 items relative to their hypothesised factor definitions forwarded by Clough *et al.* (2007). Scrutineers exhibited a comprehensive understanding of the scale development process and/or of the MT literature. The item validation task required scrutineers to read the subscale definitions and rate the extent to which each item reflected the definition on a 5-point Likert scale ranging from (1) "very poor match" to (5) "excellent match". In addition, they were asked to refine the items and to provide alternative items as they deemed appropriate. Scrutineers were then asked to review the adequacy of the proposed stem. However, no comments regarding the adequacy of the stem were made.

The ratings provided by the scrutineers were used to calculate the CVI (Lynn, 1986) for each item. The CVI is used to inform decisions regarding item removal, refinement, and preservation and has been frequently adopted in recent scale development studies in sport and exercise psychology (e.g., Bartholomew *et al.*, 2010; Williams & Cumming, 2011). The CVI was calculated by dividing the number of scrutineers who gave a rating of 4 or 5 (i.e., rated the item to be a good match or excellent match to the hypothesised definition, respectively) by the number of scrutineers. Lynn (1986) suggests that when a minimum of six scrutineers are used, CVI's close to 0.80 are acceptable. 15 items displayed CVI's of 0.67 or below and were therefore deemed invalid and were subsequently removed. Based upon the low alternative item suggestion rate (n = 13), no items were revised. All remaining items exhibited CVI's ranging from 0.83 (5/6) to 1.00 (6/6) and were thus retained. Examination of item content was conducted to ensure that the core facets of the respective factor definitions were fully represented (MacKenzie *et al.*, 2011). Analysis of item face validity revealed good content coverage with respect to the factor definitions forwarded by Clough *et al.* (2007). This resulted in a revised pool of 45 items that were subsumed to accurately reflect the 4/6C's model of MT. These items resulted in the formation of the UCMTQ.

Construction of UCMTQ document

One limitation of the scale development literature (including scale development manuals) appears to be the distinct lack of articulation of the procedures administered between the item validation phase and the construction of an operational questionnaire. Specifically, item order does not appear to be discussed nor underpinned by a rationale. This consideration would appear to be worthy of attention since it has been suggested that measurement error covariances which represent respondent characteristics may reflect bias such as yea-saying or nay-saying (response acquiescence) and social desirability (Byrne, 2010). Response acquiescence is of particular interest since the error associated with response acquiescence bias may potentially be negated when item order is randomised with respect to a) the factor to which an item is hypothesised to measure, and b), its orientation (positive or negative wording). For example, one could argue that the likelihood of observing response acquiescence would be lower when items are randomly positioned based upon their hypothesised factor and orientation. The random positioning of items would require respondents to continually attend and engage with the ongoing list of items. On the other hand, response acquiescence may be more prevalent when items are allocated (randomly or systematically) proximally relative to their hypothesised factor, or when the ratio of item orientation is biased or unbalanced (i.e., blocks of either negative or positively worded items) which may demand less attention and engagement from the respondent.

In an effort to negate the potential effects of order upon the factorial properties of the UCMTQ, a series of steps were undertaken to randomise the order of the items. First, items were assigned a number between 1 and 45. The first item from each respective factor was assigned to block 1. This process was repeated for the second item for each respective factor and so forth which resulted in 6 blocks of 6 items which were independent of their hypothesised factor. Second, each block was then submitted to a random number generator to establish the item order for each block. Once an item order was identified that did not contain an item orientation bias (i.e., multiple negative items or positive items in a row), the order of that block was established. Finally, the block numbers were submitted to a random number generator to determine their order in the UCMTQ. This process ensured that items were equally distributed in the questionnaire relative to their hypothesised factor and orientation.

Measures

University of Chichester Mental Toughness Questionnaire (UCMTQ)

The UCMTQ is a 45-item inventory which requires respondents to rate their agreement to statements on a 5-point Likert scale ranging from (1) strongly disagree to (5) strongly agree (see Appendix 2.1). The UCMTQ measures six subscales of challenge (6 items), commitment (9 items), control – emotion (7 items), control – life (8 items), confidence – abilities (7 items), and confidence – interpersonal (8 items). Example items include “I thrive in continually changing environments” (challenge); “I find myself giving up on things when the going gets really tough” (commitment); “I find it difficult to control my emotions in high pressure situations” (control - emotion); “I am in control of my life” (control - life); “I am confident in my own abilities” (confidence - abilities); “I get intimidated in social situations” (confidence - interpersonal). The UCMTQ contains 20 reversed items which are recoded prior to calculation of average scores for each subscale. Items for each factor are totalled and averaged to give a score from 1 to 5 (5 indicating the highest score on the scale). Total mean MT is calculated by totalling the scores from all 45 items and dividing by 45 (see Appendix 2.2 for UCMTQ scoring key).

Social Desirability Questionnaire (SDS)

See Study 1 for details of the SDS.

Procedure

The procedure used for Study 2 was identical to Study 1. Participants were recruited over a period of one month. Participant information sheet, consent form, demographic questionnaire, and written debrief are presented in Appendix 2.3.

Data analyses

Data screening

Data screening analyses used for Study 2 were identical to those used Study 1. Listwise deletion (n = 24) was employed where missing values were identified (Tabachnick & Fidell, 2001). Analyses identified that the data were skewed and kurtotic (see Table 4.1). Ratios below -2 and above 2 are indicative of univariate nonnormality (Fallowfield *et al.*, 2005). Analysis of multivariate normality revealed a kurtosis C. R. value of 44.419 which is above Bentler’s (2005) threshold of 5.00 and is therefore indicative of nonnormal data. Data were screened for univariate outliers using box-plots (see Appendix 2.4). A total of 129 extreme outliers (relating to 60 participants) were identified. In order to assess the impact of the extreme outliers upon the distribution of the overall data set, two independent data sets were formulated. Data set A (unscreened) comprised of the original data set (*n* = 481) and data set B (screened) comprised of the original data set minus the extreme outliers (*n* = 421). Multivariate outliers were assessed through the computation of D^2 values for each participant (see Appendix 2.5). Examination of Mahalanobis distances identified numerous multivariate outliers. Analysis of the distribution of data set A and data set B did not result in any discernible differences in the distribution of normality (see Table 4.1 & Appendix 2.6, respectively). Consequently, in light of sample preservation, those cases (*n* = 60) which were identified as univariate outliers were reinstated. Therefore, all analyses were conducted on the original data set (*n* = 481). Descriptive statistics by factor are depicted in Table 4.1. Data screening was conducted using PASW statistics 18.0.

Confirmatory Factor Analysis

The CFA procedures used to assess the factorial validity of the first- and second-order models of the four and six factor models of the UCMTQ were identical to those used in Study 1. Path diagrams for the UCMTQ are displayed in Appendix 2.7.

4.3 Results

Descriptive statistics

Descriptive statistics of UCMTQ scores are depicted in Table 4.1. Analysis of skewness and kurtosis ratios revealed that the data were nonnormally distributed. Inter-factor correlations of the UCMTQ were weak to moderate suggesting that the factors represent related yet independent components of the 4/6C's model of MT (see Table 4.2). The correlations between the respective UCMTQ factors and the SDS were generally weak which suggests that socially desirable responding had little impact on the regression weights observed (see Table 4.2). The respective factors of the UCMTQ demonstrated adequate internal reliability (α > 0.70: Nunnally & Bernstein, 1994; commitment: 0.77; control: 0.71; control – emotion: 0.70; confidence: 0.82; confidence – abilities: 0.80) except for the challenge, control – life, and confidence interpersonal subscales (α = 0.63, 0.56, 0.68, respectively). Overall, mean total MT scores demonstrated high internal reliability (α = 0.89; see Table 4.2).

Confirmatory Factor Analyses

Goodness of fit

The CFA's conducted on the hypothesised models revealed poor model fit relative to the conservative and liberal guidelines (see Study 1 Method). Table 4.3 indicates that although the first-order models and six factor models provided marginally better model fit than the second-order models and four factor models, in general, the respective fit indices are below the conservative thresholds (identified in parentheses) and the liberal thresholds.

Table 4.1. Descriptive statistics of the UCMTQ by factor.

Factors		Item	Mean	*S.D.*	S ratio	K ratio
Challenge	UCMTQ3R	I find that challenges stop me reaching my goals	3.38	0.89	-3.42	-2.86
	UCMTQ7	For me, an activity is only worthwhile doing if it stretches me to my limits	2.91	0.97	1.86	-2.80
	UCMTQ11	I see testing situations as opportunities for me to develop as a person	3.86	0.71	-7.73	6.10
	UCMTQ27	I thrive on pushing myself to the limits of my abilities	3.80	0.81	-5.56	2.21
	UCMTQ32	I get most enjoyment out of putting myself in challenging situations	3.55	0.83	-4.54	-0.38
	UCMTQ40	I thrive in continually changing environments	3.28	0.88	-2.59	-0.29
Commitment	UCMTQ8	If something is worth doing, I devote all my efforts to see it through	4.03	0.77	-6.84	4.16
	UCMTQ13	I remain committed to my goals no matter what obstacles are put in front of me	3.73	0.79	-6.16	2.22
	UCMTQ20R	I find it difficult to stay dedicated to a task when I have a tough deadline	3.26	1.07	-3.10	-4.19
	UCMTQ25R	I often find myself giving up on a task	3.92	0.80	-6.68	2.71
	UCMTQ30	I am fully committed to achieving the goals I have set myself	3.90	0.70	-6.27	5.29
	UCMTQ35R	I am unable to bounce back following failures	3.79	0.94	-6.40	0.19
	UCMTQ38R	I find myself giving up on things when the going gets really tough	3.73	0.85	-6.32	1.18
	UCMTQ41	No matter how hard things get, I see a task through to the end	3.74	0.76	-5.28	1.59
	UCMTQ44	I make myself do things whether I want to or not	3.51	0.85	-5.39	-0.47
Control - emotion	UCMTQ4	I am able to control the impact that my nerves have on me	3.29	0.91	-3.81	-2.74
	UCMTQ10R	I find it difficult to control my emotions in high pressure situations	3.19	1.06	-2.18	-4.34
	UCMTQ12R	I let my anxieties get the better of me	3.36	1.00	-3.05	-2.70
	UCMTQ16R	My emotions get the better of me when I want to do well	3.27	1.00	-2.76	-3.58
	UCMTQ23	I often hide my emotions from others	3.70	1.04	-6.08	-1.19
	UCMTQ33R	I get frustrated when things do not go my way	2.54	1.00	5.14	-1.93
	UCMTQ43	I can remain calm even in the most difficult situations	3.30	0.98	-2.71	-3.44

Table 4.1. *(continued)*

Factor		Item	Mean	*S.D.*	S ratio	K ratio
Control - life	UCMTQ2	I am in control of my life	4.04	0.71	-5.92	3.99
	UCMTQ9R	I believe that whatever is going to happen is going to happen	2.56	1.16	3.75	-3.66
	UCMTQ14	I am the major cause of my own destiny	3.98	0.80	-5.87	2.23
	UCMTQ17	I believe that I am in control of the plans I make	3.96	0.69	-9.08	11.78
	UCMTQ21R	I believe that what will be will be	2.63	1.09	3.91	-2.49
	UCMTQ31	I believe that getting what you want from life has little to do with luck	3.31	0.98	-2.19	-2.85
	UCMTQ34	I believe that what happens to me is down to my own actions	3.87	0.70	-4.70	2.59
Confidence - ability	UCMTQ1	I am confident in my own abilities	3.90	0.80	-7.89	3.26
	UCMTQ5R	I doubt myself when I have a difficult task to achieve	3.04	1.03	0.85	-4.35
	UCMTQ15R	I doubt myself when faced with setbacks	3.15	1.00	-0.66	-4.79
	UCMTQ19	My self-belief enables me to achieve my goals	3.78	0.71	-8.81	5.94
	UCMTQ24	I expect to succeed in performing important tasks	3.83	0.71	-8.72	7.90
	UCMTQ28R	When things go wrong my head can drop	2.74	1.04	4.03	-3.19
	UCMTQ36R	I lack self-belief	3.50	1.11	-4.56	-2.86
Confidence - interpersonal	UCMTQ6R	I feel that people don’t listen to what I have to say	3.36	0.90	-3.60	-1.67
	UCMTQ18	I find it easy to meet new people	3.84	0.96	-7.35	1.60
	UCMTQ22	I am an assertive person	3.55	0.83	-6.21	1.63
	UCMTQ26	I think that most people I know like me	3.81	0.72	-9.61	10.23
	UCMTQ29R	I feel that I have little control over the direction my life is taking	3.87	0.88	-6.84	2.11
	UCMTQ37	I am a good person to be around	4.00	0.62	-3.36	3.88
	UCMTQ42R	I get intimidated in social situations	3.52	1.06	-3.54	-2.89
	UCMTQ45R	I am not able to deal with awkward people	3.41	1.00	-4.68	-2.19

Note: S ratio = Skewness ratio; K ratio = Kurtosis ratio

Table 4.2. Correlations between factors of the UCMTQ and SDS scores and Cronbach's Alpha reliability coefficients.

	MT factors								
	CHAL	COM	CONT	CE	CL	CONFID	CA	CI	MTMT
CHAL	(0.63)								
COM	0.54**	(0.77)							
CONT	0.29**	0.45**	(0.71)						
CE	0.22**	0.36**	0.84**	(0.70)					
CL	0.27**	0.38**	0.76**	0.33**	(0.56)				
CONFID	0.44**	0.52**	0.58**	0.53**	0.42**	(0.82)			
CA	0.43**	0.51**	0.63**	0.60**	0.43**	0.88**	(0.80)		
CI	0.31**	0.40**	0.36**	0.30**	0.30**	0.83**	0.47**	(0.68)	
MTMT	0.62**	0.75**	0.79**	0.67**	0.62**	0.87**	0.83**	0.65**	(0.89)
SDS	0.17**	0.30**	0.21**	0.27**	0.07	0.11*	0.13**	0.04	0.24**

Note: **Significance is at $p < 0.01$. Cronbach's Alpha Reliability Coefficients are displayed in parentheses. CHAL = Challenge; COM = Commitment; CONT = Control; CE = Control - emotion; CL = Control - life; CONFID = Confidence; CA = Confidence - ability; CI = Confidence - interpersonal; MTMT = Mean Total MT; SDS = Social Desirability Scale.

Table 4.3. Summary of fit indices across hypothesised model specification.

	Fit indices					
CFA model Criterion values	CMIN/DF (< 2.00)	*P* (> 0.05)	CFI (> 0.95)	PCFI (> 0.60)	RMSEA (90% CI) (< 0.05)	AIC (lower = better)
First-order four factor	3.115	0.000	0.681	0.646*	0.066 (0.064, 0.069)	3206.887
First-order six factor	2.830	0.000	0.727	0.683*	0.062 (0.059, 0.065)	2931.694
Second-order four factor	3.238	0.000	0.622	0.629*	0.068 (0.066, 0.071)	3325.418
Second-order six factor	2.961	0.000	0.704	0.668*	0.064 (0.061, 0.067)	3062.289

Note: * denotes good fit.

Regression weights

Four factor models

The second-order four factor model revealed that the confidence factor was found to have the strongest relationship with the second-order MT factor and challenge was found to have the weakest relationship with the second-order MT factor. In line with the guidelines proposed by Comrey and Lee (1992), all four factors could be considered excellent in terms of their relationship with the second-order MT construct. Items relating to the commitment factor were found to have the strongest relationships with their hypothesised factor and items relating to the confidence factor were found to have the weakest relationships with their hypothesised factor. Using the guidelines proposed by Comrey and Lee, inspection of factor estimates shows that only 21 out of 45 items (47%) could be considered to have fair or above relationships with their hypothesised factors (see Appendix 2.8).

Estimates obtained from individual items of the challenge and commitment factors were generally considered to have fair or above relationships with their hypothesised factors, with 4 out of 6 items (66.6%) and 5 out of 9 items (55.5%) considered fair or above, respectively. Moreover, one item from the challenge factor and five items from the commitment factor were considered to have very good or above relationships with their hypothesised factors. A large proportion of estimates from the individual items of the control and confidence factors were considered to have poor relationships with their hypothesised factors, with 8 out of 15 items (53.3%) and 10 out of 15 items (66.6%) considered poor or below, respectively. However, two items from the control factor (UCMTQ4, UCMTQ12R) and four items from the confidence factor (UCMTQ1, UCMTQ5R, UCMTQ15R, UCMTQ36R) were considered to have very good or above relationships with their hypothesised factor. Inspection of estimates from the control and confidence factors indicates that while there appears to be a large proportion of items that could be considered to have poor or below relationships with their hypothesised factors, there are a small number of items which exhibit strong relationships with their hypothesised factors. Only item MTQ9R, MTQ21R, MTQ23, and MTQ31 failed to display statistically significant relationships ($p < 0.05$) with their hypothesised factors (see Appendix 2.8).

Six factor models

The second-order six factor model revealed that the confidence - abilities factor was found to have the strongest relationship with the second-order MT factor and confidence - interpersonal was found to have the weakest relationship with the second-order MT factor (see Appendix 2.8). In line with the guidelines proposed by Comrey and Lee (1992), 5 out of 6 factors could be considered excellent in terms of their relationship with the second-order MT construct, while the confidence - interpersonal factor could be classified as having a very good relationship with the second-order MT construct. Regression weights of the first- and second-order six factor models revealed that in general, items relating to the confidence - abilities factor were found to have the strongest relationships with their hypothesised factor and that items relating to the confidence - interpersonal factor were found to have the weakest relationships with their hypothesised factor. Using the guidelines proposed by Comrey and Lee (1992), inspection of factor estimates for the first- and second-order six factor models shows that only 27 out of 45 items (60%) and 26 out of 45 items (57%) could be considered to have fair or above relationships with their hypothesised factors, respectively (see Appendix 2.8).

Estimates obtained from individual items of the challenge, commitment, control – emotion, and confidence - ability factors were generally considered to have fair or above relationships with their hypothesised factors, with 4 out of 6 items (66.6%), 6 out of 9 items (55.5%), 5 out of 7 items (71.4%), and 5 out of 7 items (71.4%) considered fair or above, respectively. There was a marginal difference in estimates obtained from the second-order six factor model in that the commitment factor exhibited 5 out of 9 (55.5%) items which could be considered to have fair or above relationships with its hypothesised factors. Moreover, 1 out of 6 (16.6%), 5 out of 9 (55.9%), 4 out of 7 (57.1%), and 4 out of 7 (57.1%) items respectively, were considered to have very good or above relationships with their hypothesised factors. There was a marginal difference in the estimates obtained from the second-order six factor model in that the challenge factor exhibited 2 out of 6 (33.3%) items which could be considered to have very good to excellent relationships with its hypothesised factors (see Appendix 2.8).

A large proportion of estimates from the individual items of the control - life and confidence - interpersonal factors were considered to have poor relationships with their hypothesised factors, with 4 out of 8 items (50%) and 5 out of 8 items (62.5%) considered poor or below, respectively. Only one item from the control – life factor and one item from the confidence – interpersonal factor were considered to have very good or above relationships with their hypothesised factor. Marginal differences in estimates were obtained from items of the first-order six factor model in that the control – life factor exhibited no items that could be considered to have very good or above relationship with its hypothesised factor. Inspection of item estimates from the control - life and confidence - interpersonal factors suggests that while there appears to be a large proportion of items that could be considered to have poor or below relationships with their hypothesised factors, there are a small number of items (i.e., control - life: UCMTQ2; confidence – interpersonal: UCMTQ18, UCMTQ42R) which exhibit good or above relationships with their hypothesised factors. Only items MTQ9R, MTQ21R, and MTQ23 failed to display statistically significant relationships ($p < 0.05$) with their hypothesised factors (see Appendix 2.8).

Model re-specification

In order to fully examine the 4/6C's model of MT, independent model re-specification protocols were used in an effort to improve model fit of the hypothesised models (see Appendix 2.9 for the fit indices of the respective re-specified models). Table 4.4 summarises the fit indices of the best fitting revised models. Despite the first-order four factor revised model showing the highest level of change in CFI (ΔCFI) from its respective hypothesised model specification and providing the best fitting model, the improvements in model fit failed to provide an adequately fitting model in accordance with the conservative thresholds (identified in parentheses) and the liberal thresholds (see Study 1 Method).

Table 4.4. Summary of fit indices across best fitting revised model specification.

	Fit indices						
CFA model Criterion values	CMIN/DF (< 2.00)	*P* (> 0.05)	CFI (> 0.95)	Δ CFI	PCFI (> 0.60)	RMSEA (90% CI) (< 0.05)	AIC (lower = better)
First-order four factor	2.564	0.000	0.885	0.204	0.798*	0.057 (0.052, 0.062)	919.234
First-order six factor	2.738	0.000	0.860	0.133	0.767*	0.060 (0.056, 0.065)	1195.333
Second-order four factor	2.942	0.000	0.857	0.235	0.778*	0.064 (0.059, 0.069)	1031.807
Second-order six factor	2.917	0.000	0.825	0.121	0.760*	0.063 (0.059, 0.067)	1446.340

Note: * denotes good fit.

4.4 Discussion

The first aim of Study 2 was to extend the findings of Study 1 by regenerating items which better represent the 4/6C's model of MT. Inspection of item CVI's revealed that 45 (75%) out of the 60 items initially generated exceeded Lynn's (1986) recommended cut-off (> 0.80 when using 6 scrutineers) for item acceptance. Furthermore, 15 out of the 45 items (33%) accepted were found to have CVI's of 1 meaning that all six scrutineers rated the item as an excellent match to their hypothesised factor definition. Analysis of item CVI's appears to suggest that the items regenerated in this study were accurate representations of the definitions forwarded by Clough *et al.'s* (2007) to underpin the 4/6C's model of MT. The adoption of a rigorous item generation and evaluation protocol therefore satisfied the objectives of the item generation process outlined by MacKenzie *et al.* (2011), whereby an item pool should accurately encapsulate the core facets of the focal construct.

The second aim of Study 2 was to examine the factorial validity of the UCMTQ. The findings from the respective CFA's provided little support for the hypothesis that first- and second-order and four and six factor models of the UCMTQ would provide adequate fitting models (see Table 4.3). Although the respective models of the UCMTQ provided better model fit when compared to the respective models of the MTQ48 (see Table 4.3 & Table 3.3, respectively), the inadequate fit indices obtained in Study 2 indicate that the UCMTQ is not a valid measure of the 4/6C's model of MT which it intends to capture. Inspection of item regression weights provided support for the poor model fit of the UCMTQ in that the relationships between the items and their respective factors were generally weak. For the four factor models (first- and second-order), findings revealed that only 21 out of 45 items (47%) could be considered to have fair or above relationships with their hypothesised factors. For the first-order six factor model and second-order six factor model only 27 out of 45 items (60%) and 26 out of 45 items (57%) could be considered to have fair or above relationships with their hypothesised factors, respectively.

The findings from Study 2 are therefore comparable to those observed for the four and six factor models examined in Study 1 whereby 24 out of 48 items (50%) and 32 out of 48 items (66.6%) could be considered to have fair or above relationships with their hypothesised factors, respectively. This finding shows that only approximately half of the items examined in Study 1 and Study 2 satisfied the minimum cut-off of 0.45 for estimate classification (Comrey & Lee, 1992), which ultimately indicates poor item representation. Findings also revealed a number of items which had particularly low relationships with their hypothesised factors. For the four factor models, one item from the challenge factor (UCMTQ7), one item from the commitment factor (UCMTQ44), six items from the control factor (UCMTQ9R, UCMTQ21R, UCMTQ23, UCMTQ31, UCMTQ33R, UCMTQ34), and two items from the confidence factor (UCMTQ37, UCMTQ45R) were considered to be too low for interpretation. For the six factor models, one item from the challenge factor (UCMTQ7), two items from the control – emotion factor (UCMTQ23, UCMTQ33R), and three items from the control – life factor (UCMTQ9R, UCMTQ21R, UCMTQ31) were considered to be too low for interpretation. These findings add to findings of Study 1 and Gucciardi *et al.* (2012) whereby items relating to the control and confidence factors were seen to be the major contributors to the poor fitting models examined.

Similarly, item UCMTQ23 **-** "I often hide my emotions from others" demonstrated a negative relationship with its hypothesised control – emotion factor which is also consistent with the findings of Study 1 and the research by Crust and Swann (2011b) and Gucciardi *et al.* (2012). Given that item UCMTQ23 represents the same conceptual space of the control – emotion factor (i.e., "less likely to reveal their emotional state to other people"; Clough *et al.,* 2007, p. 4) as item MTQ34, the observed negative relationship could be attributed to the conceptual content of the control – emotion factor. Specifically, research examining emotional control (e.g., Nicholls & Polman, 2007) has emphasised the importance of being able to effectively control one's emotions to maintain/enhance performance as opposed to supporting the use of covert coping strategies. This apparent disparity appears to suggest that Clough *et al.* may not have accurately defined and adequately represented the conceptual space of their control – emotion factor.

Despite Study 2 providing little support for the factorial validity of the UCMTQ, findings were consistent with the studies hypotheses and the findings of Study 1 whereby the first-order models and six factor models provided marginally superior fitting solutions when compared to second-order and four factor models, respectively (see Table 3.3 & Table 4.3 for Study 1 and Study 2 fit indices, respectively). This finding appears to suggest that the UCMTQ is best understood when underpinned by Earle's (2006) 6C's model of MT and when the latent constructs are specified to be correlated. Further analysis of the UCMTQ's factorial validity revealed that although model re-specification protocols improved model fit across all models tested, these improvements failed to provide an adequately fitting model in accordance with the conservative and liberal thresholds (see Table 4.4 & Appendix 2.9). This finding provides further evidence to indicate that the UCMTQ is not a valid measure of the 4/6C's model of MT. The findings from Study 2 therefore suggest that although the items of the UCMTQ were evaluated and deemed adequate representations of their hypothesised factors by six expert scrutineers, the items generally appear to be poor indicators when analysed collectively using CFA. This would suggest that the items constructed in Study 2 are comparably inadequate to those examined in Study 1.

Given that analysis of UCMTQ item content provided support for their representation of the 4/6C's model of MT, the inadequate regression weights and fit indices obtained in Study 2 may be attributed to the structure and content of the factor definitions proposed by Clough *et al.* (2007). As previously discussed in Chapter 2, close inspection of the information provided by Clough *et al.* reveals that only the characteristics thought to make-up each construct were detailed, as opposed to definitive definitions to represent each construct (see Table 2.1). Given the need to collect a representative set of definitions from the literature to deduce and construct accurate definitions (Sartori, 1984), the need to specify the conceptual domain and entity of the construct (MacKenzie *et al.*, 2011), and the need for these steps to inform the generation of clear and concise construct definitions (Churchill, 1979; Hinkin, 1995), it could be argued that Clough *et al.'s* articulation of only the characteristics purported to make-up the 4/6C's model of MT is insufficient to satisfy the aforementioned requirements when developing a valid measurement tool. This in turn, constitutes a major limitation in the development of the definitions to underpin the 4/6C's model of MT, and ultimately the underpinning framework for the MTQ48.

There are also problems with the content of the information forwarded by Clough *et al.* (2007). Despite Clough *et al.'s* (2002) and Earle's (2006) 4/6C's model of MT being embedded in the 3C's of hardiness, close inspection of the characteristics forwarded by Clough *et al.* (2007) appear to be partially different to the characteristics identified in the hardiness literature. Table 4.5 summarises the key themes in the respective definitions detailed by Clough *et al.* and the hardiness literature. Clough *et al.* state that challenge "describes the extent to which individuals see challenges as opportunities" (p.4) and that individuals high in challenge "will actively seek them out and will identify problems as ways for self-development" (p.4). Clough *et al.* also state that "at the other end challenges are perceived as problems and threats" (p.4). The characteristics detailed by Clough *et al.* (see Table 2.1) therefore show some parallels to the hardiness literature whereby individuals high in challenge are thought to seek out challenges (Kobasa *et al.*, 1982) and that change is perceived as a means for personal development rather than threats to security (Berlyne, 1964; Csikzentmihalyi, 1975; Maddi *et al.*, 1965).

Despite these similarities, there are a number of facets relating to challenge which Clough *et al.* (2007) do not acknowledge including the notion that change is a normal part of life (Berlyne, 1964; Csikzentmihalyi, 1975; Maddi *et al.* 1965), that individuals high in challenge are cognitively flexible and know how to cope by using previous experiences (Kobasa, 1979). Clough *et al.* state that highly committed individuals "carry out tasks successfully despite any problems or obstacles that arise whilst achieving their goal" (p.5) and are "able to handle and achieve things to tough unyielding deadlines" (p. 5). Although Clough *et al.* appear to partially capture the characteristic of not giving up under pressure easily (Judkins, Arris, & Kenner, 2005), there are numerous characteristics which have not been captured. Specifically, involving

Table 4.5. Summary of characteristics of the respective literature.

	4/6C's of MT (Clough *et al.* 2007)	**3C's of Hardiness**
Challenge	Challenges are opportunities Seek out challenges Identify problems as ways for self-development Individuals low in challenge perceive new environments as threatening	Change is normal[1, 2, 3] Seek out challenges[4] Changes are incentives for personal development rather than threats to security[1, 2, 3] Know how to cope by using previous experiences[5] Cognitively flexible[5]
Commitment	Perform tasks successfully despite obstacles	Involve oneself in whatever one is doing[6] Find an interest in whatever one is doing[7] Stress mitigated by a sense of purpose[5] Involvement with others as a coping mechanism[8] Won't give up attitude[9]
Control	Control over work Confident in multitasking Ability to control emotions Less likely to reveal their emotional state to others Control over one's life Control over plans	Feel and act as if one is being influential[10, 11, 12] Enhanced stress resistance[4] Dealing with stress as part of a life plan[4] Broad range of coping strategies to draw upon in stressful situations[4] Continue to be influential even in stressful situations[13]

Primary sources. [1]Berlyne, 1964; [2] Csikzentmihalyi, 1975; [3]Maddi *et al.*, 1965; [4] Kobasa *et al.*, 1982; [5]Kobasa, 1979; [6] Maddi *et al.*,1982; [7]Maddi & Hightower, 1999; [8] Antonovsky, 1974; [9] Judkins *et al.*, 2005; [10] Averill, 1973; [11]Phares, 1976; [12]Seligman, 1975; [13]Maddi, 2008.

oneself in (Maddi *et al.*, 1982) and finding an interest in whatever one is doing (Maddi & Hightower, 1999) and stress being mitigated by a sense of purpose (Kobasa, 1979) and through involvement with others (Antonovsky, 1974) are not captured.

Clough *et al.'s* (2007) decision to split the control construct makes it difficult to evaluate its content relative to the hardiness literature. Notwithstanding, Clough *et al.* state that control – emotion describes individuals who "are better able to control their emotions" (p. 4) and are "less likely to reveal their emotional state to other people" (p. 4). In regards to control – life, Clough *et al.* state that individuals high in this construct "are more likely to believe that they control their lives" (p. 4) and "feel that their plans will not be thwarted and that they can make a difference" (p. 4). The characteristics forwarded by Clough *et al.* therefore show some general similarities with the hardiness literature whereby individuals high in control are thought to have an enhanced stress resistance and have a broad range of coping strategies to draw upon in stressful situations (e.g., Kobasa, 1982). However, Clough *et al.'s* characteristics do not explicitly encapsulate the notion that individuals high in control feel and act as if they are influential over their lives (Averill, 1973; Phares, 1976; Seligman, 1975), deal with stress as part of a life plan (Kobasa *et al.*, 1982), and continue to be influential even in stressful situations (Maddi, 2008). Clough *et al.* also offer characteristics which are presented under a general control construct (see Table 2.1). Clough *et al.* state that individuals high in control "feel that they are in control of their work and of the environment in which they work" (p. 4) and are "capable of exerting more influence on their working environment and are more confident about working in complex or multi-tasked situations" (p. 4). Despite Clough *et al.'s* characteristics being somewhat captured by the notion of being influential (Averill; Phares; Seligman), the other descriptors appear to be redundant in regards to their ability to capture the characteristics outlined in the hardiness literature.

One may also question the characteristics forwarded by Clough *et al.* (2007) to underpin the confidence construct. Given the limited research examining the concept of dispositional or global self-confidence and that self-confidence is often used without a theoretical basis (Bandura, 1997), the concept of self-esteem offers a related and theoretically embedded alternative to the examination of dispositional self-confidence. Unlike self-confidence, self-esteem is thought to have a trait component and has been argued to be made up of two interrelated components; competence and self-worth (e.g., Branden, 1969). Using self-esteem as a conceptual basis, one may tentatively assess the conceptual coverage of Clough *et al.'s* confidence characteristics. Clough *et al.* state that individuals high in confidence abilities "are more likely to believe that they are a truly worthwhile person" (p. 5) and are "less dependent on external validation and tend to be more optimistic about life in general" (p. 5). In regards to confidence – interpersonal, Clough *et al.* state that individuals high in this construct "tend to be more assertive" (p. 5), "less intimidated in social settings and less likely to push themselves forward in groups" (p. 5), and are "better able to cope with difficult or awkward people" (p. 5).

Similar to the control construct, Clough *et al.* (2007) also present characteristics under a general confidence descriptor. Clough *et al.* state that individuals high in confidence "have the self-belief to successfully complete tasks" (p. 5), "less confident individuals are also less likely to be persistent and may make more errors" (p. 5), and individuals low in confidence "will be unsettled by setbacks and will feel undermined by these" (p.5). Inspection of Clough *et al.'s* confidence characteristics reveals that although the components of self-worth (i.e., the belief that one is a worthwhile person) and competence (i.e., individuals high in confidence have the self-belief to successfully complete tasks) are partially represented, there are a number of additional facets included by Clough *et al.* which are not explained by the tenets of self-esteem. Upon evaluating the definitions underpinning the 4/6C's model of MT, it appears that Clough *et al.* did not accurately define and adequately represent the constructs underpinning the 4/6C's model of MT. One may therefore argue that the procedures adopted by Clough *et al.* to define the constructs underpinning the 4/6C's model of MT may explain the poor factorial validity of the UCMTQ.

4.5 Summary

The aim of Study 2 was to regenerate items which better represent the 4/6C's model of MT in an effort to develop a valid measure of MT. Although analysis of UCMTQ item content provided support for their representation of the 4/6C's model of MT, the CFA findings revealed little support for the factorial validity of the hypothesised and revised models of the UCMTQ. The findings appear to suggest that Clough *et al.* (2007) did not accurately define and adequately represent the constructs underpinning the 4/6C model of MT (i.e., challenge, commitment, control, and confidence), which appear to explain the poor fitting models and poor regression weights observed in Study 2. In order to fully test the utility of the 4/6C's model of MT, Study 3 will re-develop construct definitions which more accurately represent the core traits underpinning the 4/6C's model of MT. The generation of construct definitions which accurately represent dispositional challenge, commitment, control, and confidence will afford the generation of items which accurately capture their intended construct domain, which in turn, could enable the provision of a valid instrument to measure the traits thought to underpin MT.

CHAPTER 5.0

STUDY 3: THE DEVELOPMENT AND FACTORIAL ANALYSIS OF THE HARDINESS CONFIDENCE QUESTIONNAIRE

5.1 Introduction

The findings from Study 1 provided little support for the factorial validity of the MTQ48. Close inspection of MTQ48 item content revealed that in general, the items may be inaccurate representations of the factor definitions forwarded by Clough *et al.* (2007) to underpin the 4/6C's model of MT. Study 2 regenerated items to better represent the 4/6C's model of MT in an effort to provide a valid instrument to measure the 4/6C's model of MT. Although analysis of the UCMTQ item content provided support for their representation of the 4/6C's model of MT, the findings revealed little support for the factorial validity of the developed UCMTQ. Close inspection of the definitions generated to underpin the 4/6C's model of MT revealed that Clough *et al.* did not accurately define and adequately represent the content of its focal constructs in accordance with the scale development literature (e.g., Sartori, 1984; MacKenzie *et al.*, 2011; Churchill, 1979; Hinkin, 1995; see Table 4.5), which in turn, could explain the poor factorial validity of the UCMTQ. In order to fully examine Clough and colleagues' approach to MT, research is required to develop an instrument which captures the core traits thought to underpin MT (i.e., hardiness and confidence).

Many researchers (e.g., MacKenzie, 2003; MacKenzie *et al.*, 2011; Nunnally & Bernstein, 1994) have emphasised the importance of adequately defining the construct under examination prior to constructing an instrument for its measurement. Specifically, Nunnally and Bernstein argue that developing a constructs definition is the most important aspect of the scale development process because "there is no way to know how to test the adequacy with which a construct is measured without a well specified domain" (p. 88). Failure to adequately define the construct can lead to a number of issues including confusion regarding what the construct does and does not refer to, poor representation at item level, and invalid conclusions concerning the relationships with other constructs (MacKenzie *et al.*). In light of the findings of Study 2, research is required to redevelop the 4/6C's model of MT by regenerating definitions which better represent the traits underpinning the model (hardiness and confidence). This in turn will shed light on its utility as a valid conceptualisation of MT. Specifically, research is required to develop Clough *et al.'s* (2002) 4C's model of MT as this model structure provides a more accurate representation of dispositional challenge, commitment, control, and confidence. However, given that self-esteem provides a conceptually embedded construct to examine dispositional self-confidence which encompasses both self-competence and self-worth, the development of a 5C's model is warranted.

Therefore, the first aim of Study 3 was to regenerate definitions which better represent the core traits underpinning MT (i.e., challenge, commitment, control, confidence). The second aim of Study 3 was to generate an instrument which accurately represents the factor definitions of the redeveloped 5C's model of MT, namely the Hardiness Confidence Questionnaire (HCQ). The third aim of Study 3 was to examine the factorial validity of the HCQ in an effort to develop a valid measure of the traits thought to underpin MT. EFA and CFA will be used to examine the factorial validity of the first- and second-order models of the HCQ. It is hypothesised that both the first- and second-order five factor models of the HCQ will provide adequate model fit. Given that it is theoretically plausible for hardiness to predict confidence and for confidence to predict hardiness (see Chapter 2 for a review of the theoretical relationships between hardiness and confidence), structural models were used to examine the relations among the HCQ's underlying latent factors. It was hypothesised that the structural models would offer superior model fit when compared to the first- and second-order five factor models of the HCQ.

5.2 Method

Participants

Participants were 330 competitive student athletes from a University in Southern England. Sixteen did not fully complete the HCQ and were subsequently removed from the data analysis. EFA was used to examine the factor structure of a small sub-sample (*n* = 105; exploratory sample), with the remaining participants (*n* = 209) being used for CFA (confirmatory sample).

Exploratory sample

The Kaiser-Meyer-Olkin (KMO) statistic was computed as a test for sampling adequacy prior to conducting EFA. Using a sample of 105 participants, the KMO value for the HCQ was 0.761. According to Hutcheson and Sofroniou (1999), this value can be classified as good, and it is above the recommended value of 0.60 required to conduct adequately powered EFA (Garson, 2006). Consequently, the exploratory sample comprised of approximately 33% (*n* = 105) of the overall sample pool and consisted of 67 males and 38 females (*M* age = 20.02 years, *S.D.* = 1.67). The majority of participants (93.3%) were White British, with the remaining participants (*n* = 7) consisting of a variety of ethnic backgrounds such as Black or Black British African (*n* = 2), Asian Pakistani (*n* = 1), and Mixed White and Black African (*n* = 1). Participants had a mean of 9.50 years (*S.D.* = 4.01) competitive playing experience in their primary sport. One participant did not specify the number of years competitive playing experience in their primary sport. Participants were involved in both team sports (*n* = 81), such as football (*n* = 40), netball (*n* = 10) rugby (*n* = 8), and cricket (*n* = 5), and individual sports (*n* = 24), such as athletics (*n* = 5), swimming (*n* = 5), and trampolining (*n* = 4). The highest level of primary sport participation ranged from recreational (*n* = 1) through intervarsity (*n* = 6), club (*n* = 40), county (*n* = 30), regional (*n* = 10), national (*n* = 8), and international (*n* = 10) level.

Confirmatory sample

The confirmatory sample comprised of approximately 67% (*n* = 209) of the overall sample pool. Using Stevens (1996) sample calculation as a guide (five participants per variable), the confirmatory sample appears to provide an adequately powered sample to examine the factorial validity of the 27-item HCQ (i.e., *n* > 135). The confirmatory sample consisted of 140 males and 68 females (*M* age = 19.91 years, *S.D.* = 2.83). The majority of participants (92.3%) were White British, with the remaining participants (*n* = 16) consisting of a variety of ethnic backgrounds such as White other (*n* = 4), Mixed other (*n* = 2), and any other ethnic background (*n* = 2). Participants had a mean of 9.42 years (*S.D.* = 3.63) competitive playing experience in their primary sport. One participant did not specify their gender. Participants were involved in both team sports (*n* = 140), such as football (*n* = 84), rugby (*n* = 22), cricket (*n* = 14), and netball (*n* = 10), and individual sports (*n* = 69), such as athletics (*n* = 11), swimming (*n* = 7), and tennis (*n* = 9). The highest level of primary sport participation ranged from recreational (*n* = 3) through intervarsity (*n* = 14), club (*n* = 85), county (*n* = 67), regional (*n* = 11), national (*n* = 20), and international (*n* = 9) level.

Construct refinement

Given the extensive body of literature which has examined hardiness (Maddi, 2002) and the need for the researcher to define the focal construct in a manner that is consistent with prior research (MacKenzie, 2003), an examination of the hardiness literature was undertaken to identify definitions of challenge, commitment, and control. Inspection of the definitions provided by Kobasa *et al.* (1982) to underpin the 3C's model of hardiness (Kobasa, 1979) appear to satisfy the recommended guidelines for developing a conceptual definition (see Figure 2.2 for a review of the recommended guidelines). According to Kobasa and colleagues, the "challenge disposition is expressed as the belief that change rather than stability is normal in life and that the anticipation of changes are interesting incentives to growth rather than threats to security (Kobasa *et al.,* p. 170)". The "commitment disposition is expressed as a tendency to involve oneself in (rather than experience alienation from) whatever one is doing or encounters (Kobasa *et al.,* p. 169)". The "control disposition is expressed as a tendency to feel and act as if one is influential (rather than helpless) in the face of the varied contingencies of life (Kobasa *et al.,* p. 169)". In line with MacKenzie *et al.* (2011), evaluation of Kobasa *et al.'s* definitions clearly incorporate information to specify the constructs conceptual domain (e.g., use of "disposition" and "belief" to orientate what the construct refers to) and entity (e.g., use of "one" and "oneself" to specify the object to which the property applies). Inspection of the definitions forwarded by Kobasa *et al.* also appear to satisfy the final stage of the construct definition process as reflected by the clear and concise nature of the constructed definitions (e.g., Churchill, 1979; Hinkin, 1995; MacKenzie *et al.*, 2011). Kobasa *et al.'s* (1982) definitions therefore appear to provide adequate representations of challenge, commitment, and control and were subsequently taken forward to underpin the redeveloped 5C's model of MT.

Similarly, the confidence literature was assessed in an effort to identify an adequate definition. Inspection of the self-confidence literature revealed that there is limited research examining the concept of dispositional or general self-confidence, with the majority of the literature not being underpinned by a sound theoretical framework (Bandura, 1997). Although researchers have regarded self-confidence and self-efficacy to be relatively similar, Bandura differentiated between these concepts in that self-efficacy is a specific perception about one's ability to conduct a particular behaviour. Consequently, the situation specific grounding of self-efficacy renders the concept inappropriate for the purpose of providing a definition of dispositional self-confidence. In an effort to identify an established psychological concept relating to dispositional self-confidence, inspection of the self-esteem literature was undertaken. Unlike self-efficacy, self-esteem is thought to have a trait and state component and is seen in such phrases as "trait versus state" self-esteem (Leary & Downs, 1995), "stable versus unstable" self-esteem (Greenier, Kernis, & Waschull, 1995), or "global versus situational" self-esteem (Harter, 1999). One prominent perspective of trait self-esteem defines the concept in terms of competence and self-worth (Mruk, 2006). Branden (1969) offered the first definition to conceptualise self-esteem as a combination of competence and self-worth. According to Branden:

"self-esteem has two interrelated aspects: it entails a sense of personal efficacy and a sense of personal worth. It is the integrated sum of self-confidence and self-respect. It is the conviction that one is *competent* to live and *worthy* of living" (p. 110).

This approach has been advocated by other theorists and has been described as a "dual model" of self-esteem (Franks & Marolla, 1976), a "two-factor" theory (Tafarodi & Swann, 1995), or as a "multidimensional approach" (Harter, 1999; O'Brien & Epstein, 1988).

Given the lack of research which has examined dispositional self-confidence, the lack of theoretical underpinning used to examine the construct (Bandura, 1997), and the apparent conceptual associations between self-esteem and self-confidence, Branden's (1969) definition of self-esteem appears to provide a well-grounded approach to define trait confidence. In line with the scale development literature (see Figure 2.2), the following definition of confidence was constructed; "The confidence disposition is expressed as a belief that one is competent (rather than inept) both cognitively and physically. Confident persons have a belief system that enables them to believe they have a high general self-worth". Evaluation of the constructed confidence definition supports its adequacy in capturing the construct's conceptual domain (i.e., use of "disposition" and "belief") and entity (i.e., use of "one"; MacKenzie *et al.*, 2011). Furthermore, the constructed definition appears to capture information to reflect beliefs of competence and self-worth in clear and concise language (e.g., Churchill, 1979; Hinkin, 1995; MacKenzie *et al.*). The constructed confidence definition therefore provides a conceptually grounded approach to examine dispositional confidence. Given that confidence was conceptualised as a combination of competence and self-worth (Branden, 1969), confidence was split into two independent constructs. This resulted in the conception of the revised 5C's model of MT.

Stem generation

The stem generated for the HCQ was an extension of the stem used for the UCMTQ. In an effort to further clarify how respondents should read each item, "In my general life" was used as the stem of the HCQ (see Appendix 3.1 for HCQ stem).

Item generation

The item generation process underpinning the HCQ was identical to that adopted in generating items for the UCMTQ (see Study 2). An iterative item refinement procedure resulted in 39 items (challenge: n = 14; commitment: n = 7; control: n = 9; confidence – competence: n = 6; confidence – self-worth: n = 3) being taken forward for item validation.

Item validation

An item validation task identical to that administered in Study 2 was constructed and employed. Pilot testing was administered with the same postgraduate students used in Study 2. Minor structural changes were made from piloting the item validation task. An independent group of expert scrutineers (used in Study 2: n = 3; not previously used: n = 3) were recruited to review the adequacy of the initial 39 item pool in relation to their hypothesised factor definitions. Out of the 39 items generated to represent challenge, commitment, control, confidence – competence, and confidence – self-worth, 9 out of 14, 1 out of 7, 2 out of 9, 1 out of 6, and 1 out of 3 items displayed CVI's of 0.67 (4/6) or below, respectively. Using Lynn's (1986) recommended 0.80 cut-off for item acceptance when using six scrutineers, these items were considered invalid representations of their hypothesised factors and were consequently removed. However, this resulted in the challenge factor only being represented by 5 items. In order to ensure that the challenge items adequately encapsulated the conceptual space of its hypothesised factor, one challenge item with a CVI of 0.67 ("I do not like to make changes to my everyday schedule") was reinstated. This resulted in the challenge factor being represented by 6 items. In addition, item removal also led to the confidence – self-worth factor only being represented by two items. Researchers (e.g., Clark & Watson, 1995; Kline, 2000; MacKenzie *et al.*, 2011) suggest that a minimum of three items should be generated to adequately represent a given factor. Accordingly, two additional items were generated in an effort to enhance the conceptual coverage of items purporting to measure the self-worth facet of confidence.

Based upon the feedback received from the scrutineers, some items (n = 7) were refined to improve their readability, clarity, and adequacy. Only one item representing the commitment factor (“I stay involved in whatever I am doing”) was subject to major refinement due to its perceived vagueness by one scrutineer. This item was refined to “I stay committed in the face of adversity”. One confidence - competence item (“I do not think I have what it takes to succeed”) was perceived by one scrutineer to be very ambiguous. In light of this feedback, this item was removed in an effort to maintain the clarity of items tapping into the confidence – competence construct. This resulted in the confidence - competence factor being represented by four items. No comments regarding the adequacy of the stem were made. All remaining items exhibited CVI’s ranging from 0.83 (5/6) to 1.00 (6/6) and were thus retained. Examination of item content revealed that the conceptual space of the respective factor definitions were adequately represented by the accepted item pool (MacKenzie *et al.*, 2011). This resulted in the generation of 27 items (commitment: n = 6; challenge: n = 6; control: n = 7; confidence – competence: n = 4; confidence – self-worth: n = 4) subsumed to represent the revised 5C’s model of MT. These items resulted in the formation of the HCQ.

Construction of HCQ document

The processes underpinning the construction of the HCQ document were identical to those adopted in the construction of the UCMTQ (see Study 2).

Measures

Hardiness Confidence Questionnaire (HCQ)

The HCQ is a 27-item inventory which requires respondents to rate their agreement to statements on a 5-point Likert scale ranging from (1) strongly disagree to (5) strongly agree. The HCQ measures five subscales of challenge (6 items), commitment (6 items), control (7 items), confidence – competence (4 items), and confidence – self-worth (4 items). Example items include “I believe that change enables me to grow as a person” (challenge); “I am eager to stay dedicated to a task” (commitment); “What happens to me is my own doing” (control); “I am confident in my own abilities” (confidence – competence); “I doubt whether I am a worthwhile person” (confidence – self-worth). HCQ scores were reversed accordingly (see Appendix 3.1 for HCQ). The HCQ contains 13 reversed items which are recoded prior to calculation of average scores for each subscale. Items for each factor are totalled and averaged to give a score from 1 to 5 (5 indicating the highest score on the scale). Total mean MT is calculated by totalling the scores from all 27 items and dividing by 27 (see Appendix 3.2 for HCQ scoring key).

Social Desirability Questionnaire (SDS)

See Study 1 for details of the SDS.

Procedure

The procedure used for Study 3 was identical to Study 1. Participants were recruited over a period of two weeks. Participant information sheet, consent form, demographic questionnaire, and written debrief are presented in Appendix 3.3.

Data analyses

Exploratory sample

Data screening

Data screening procedures were identical to those used in Study 1. Analyses identified that the data were marginally skewed and kurtotic (see Table 5.1). Data were screened for univariate outliers using box-plots (see Appendix 3.4). A total of 68 extreme outliers (relating to 38 participants) were identified. In order to assess the impact of the extreme outliers upon the distribution of the overall data set, two independent data sets were formulated. Exploratory data set A (unscreened) comprised of the original data set (n = 105) and exploratory data set B (screened) comprised of the original data set minus the extreme outliers (n = 67). Analysis of the distribution of data set A and data set B did not result in any discernible differences in the distribution of normality (see Table 5.1 and Appendix 3.5, respectively). Consequently, in light of sample preservation, those cases (n = 38) which were identified as univariate outliers were reinstated. Therefore, all analyses were conducted on the original data set (n = 105). Descriptive statistics of the exploratory sample by factor are depicted in Table 5.1.

Exploratory factor analyses

An EFA was conducted on the HCQ to examine the adequacy of the items to measure the hypothesised revised 5C's model of MT and to examine which items in the data set form coherent subsets that are relatively independent of one another (Tabachnick & Fidell, 2001). Given that the revised 5C's model of MT was underpinned by independent latent constructs, principle components analysis (PCA) with orthogonal (varimax) rotation was implemented (Tabachnick & Fidell). The use of EFA is advocated during the early stages of scale development to avoid mis-specification of the number of factors and to maximise the convergent and discriminant validity of the items representing each factor (Gerbing & Hamilton, 1996; Hurley *et al.*, 1997; Kelloway, 1995).

Confirmatory sample

Data screening

Data screening procedures were identical to those used in Study 1. Analyses identified that the data were marginally skewed and kurtotic (see Table 5.4). Analysis of multivariate normality revealed a kurtosis C. R. value of 27.98 which is above Bentler's (2005) threshold of 5.00 and is therefore indicative of nonnormal data. Data were screened for univariate outliers using box-plots (see Appendix 3.6). A total of 97 extreme outliers (relating to 49 participants) were identified. In order to assess the impact of the extreme outliers upon the distribution of the overall data set, two independent data sets were formulated. Confirmatory data set A (unscreened) comprised of the original data set (n = 209) and confirmatory data set B (screened) comprised of the original data set minus the extreme outliers (n = 160). Multivariate outliers were assessed through the computation of the D^2 for each participant (see Appendix 3.7). Examination of Mahalanobis distances identified numerous multivariate outliers. Analysis of the distribution of data set A and data set B did not result in any discernible differences in the distribution of normality (see Table 5.4 & Appendix 3.8, respectively). Consequently, in light of sample preservation, those cases (n = 49) which were identified as univariate outliers were reinstated. Therefore, all analyses were conducted on the original data set (n = 209). Descriptive statistics of the confirmatory sample by factor are depicted in Table 5.4.

Confirmatory Factor Analysis

CFA was used to examine and confirm the factor structure extracted in the EFA and further refine the structure of the scale if necessary (Marsh, 2007). The CFA procedures used to assess the factorial validity of the first- and second-order models of the HCQ were identical to those used in Study 1. Path diagrams for the HCQ are displayed in Appendix 3.9.

Structural Equation Modelling

Given that researchers (e.g., Gignac, 2009; Marsh *et al.,* 2010) have emphasised the importance of examining the dimensionality of an instrument's conceptual model, Study 3 used structural models to further examine the factorial validity of the HCQ (see Chapter 2 for a discussion on the HCQ's alternative model structures). Specifically, the first structural model examined the relations among the hardiness and confidence constructs whereby challenge, commitment, and control were hypothesised to predict confidence (HPC), whereas the second structural model was modelled whereby confidence was hypothesised to predict challenge, commitment, and control (CPH). The procedures used to examine the factorial validity of the structural models were identical to those used in Study 1. Path diagrams of the respective structural models are presented in Figure 5.1 and Figure 5.2. Structural models were analysed using AMOS statistics 18.0.

5.3 Results

Exploratory sample

Descriptive statistics

Descriptive statistics of HCQ scores using the exploratory sample are depicted in Table 5.1. Analysis of skewness and kurtosis ratios revealed that the data were marginally nonnormally distributed. Inter-factor correlations of the HCQ were weak to moderate suggesting that the factors represent related yet independent components of the revised 5C's model of MT (see Table 5.3). The correlations between the respective factors of the HCQ and the SDS were weak which suggests that socially desirable responding had little impact on the factor loadings observed in the EFA (see Table 5.2). One out of the five factors measured in the HCQ demonstrated adequate internal reliability ($\alpha > 0.70$: Nunnally & Bernstein, 1994; confidence – self-worth). Overall, the HCQ demonstrated high internal reliability ($\alpha = 0.88$). Table 5.2 displays the respective Cronbach's Alpha coefficients.

Exploratory factor analyses

Following analysis of sample adequacy, bivariate correlations were conducted and identified no evidence for multicollinearity as all correlations were below 0.80 (Stevens, 1996; see Table 5.2). PCA using orthogonal (varimax) rotation identified eight factors with eigenvalues greater than 1.0 (see Appendix 3.10). However, the use of PA has been proposed as a more reliable and accurate method of identifying the number of factors within a data set (Franklin *et al.,* 1995). In PA, the eigenvalues derived from research data are compared to those from a random sample matrix of identical dimensionality to the research data set (i.e., identical sample size). Component PCA eigenvalues which are greater than their respective component PA eigenvalues derived from the random data set are retained (Franklin *et al.*). Following the use of PA, a three factor model was suggested (see Appendix 3.11). However, given that the eigenvalue for the fourth component (see Appendix 3.10) was only marginally lower than the estimated Monte Carlo average (see Appendix 3.11) and that the scree plot suggested a four factor model (see Appendix 3.12), four factors were extracted. Researchers (e.g., Gerbing & Hamilton, 1996; Tabachnick & Fidell, 2001) have emphasised the need to use multiple indicators to inform the extraction of factors in EFA. This model explained approximately 49.7% of the cumulative variance. However, inspection of the factors, associated variables, and rotated factor loadings provided an unclear solution (see Appendix 3.13 for rotated factor loadings).

In an effort to enhance the clarity of the identified four factor solution, a secondary EFA was conducted whereby the amount of factors to be extracted was restricted to four. Although PCA using orthogonal (varimax) rotation identified eight factors with eigenvalues greater than 1.0 (see Appendix 3.14 for eigenvalues), PA and examination of the scree plot supported the extraction of a four factor model (see Appendix 3.15 for scree plot). The secondary four factor model explained approximately 49.7% of the cumulative variance. Defining variables of each factor were characterised as those with factor loadings above 0.40 which did not possess any cross-loadings or mis-loadings with an unhypothesised component (Garson, 2006). The factors, associated variables and rotated factor loadings of the secondary four factor model are depicted in Table 5.3.

Table 5.1. Descriptive statistics of the HCQ by factor using the exploratory sample.

Factors		Item	Mean	*S.D.*	S ratio	K ratio
Challenge	HCQ3R	I do not look forward to challenging situations	3.55	1.02	-2.38	-0.65
	HCQ10	I believe that change enables me to grow as a person	3.96	0.72	-5.12	6.93
	HCQ13	I see challenges in my life as opportunities for me to develop as a person	4.03	0.79	-2.75	0.58
	HCQ20	Changes in my daily routine encourage me to learn	3.52	0.76	-2.33	-0.51
	HCQ24R	I do not like to make changes to my everyday schedule	3.06	0.93	-0.18	-0.83
	HCQ26	I believe that change is a normal part of life	4.17	0.71	-5.24	8.04
Commitment	HCQ2	I am eager to stay dedicated to a task	3.95	0.67	-3.90	4.28
	HCQ6R	I often find myself disengaging from tasks	3.15	0.99	-0.29	-1.72
	HCQ8R	I find it difficult to stay committed to whatever I am doing	3.86	0.95	-3.20	0.24
	HCQ11R	I find it difficult to stay committed to achieving my goals	3.70	0.98	-3.58	0.76
	HCQ18	I stay committed to tasks even in the face of difficulty	3.77	0.79	-3.79	2.73
	HCQ23R	I lack commitment	4.05	0.90	-4.47	2.15
Control	HCQ4	I can influence the path that my life takes	4.11	0.81	-5.56	6.68
	HCQ12	I am an influential person	3.60	0.75	-1.19	-0.36
	HCQ14R	Events in my life are determined by others	3.27	1.06	-0.67	-1.64
	HCQ16	Events in my life are shaped by my own actions	4.01	0.74	-1.91	0.19
	HCQ19R	My actions have little influence on my life	4.08	0.98	-6.10	4.62
	HCQ22	What happens to me is my own doing	3.76	0.96	-3.32	1.27
	HCQ27R	I feel that I have little control over the direction my life is taking	3.95	0.97	-2.83	-1.06
Confidence - competence	HCQ1	I am confident in my own abilities	3.91	0.75	-7.74	10.48
	HCQ7	I expect to succeed in performing important tasks	3.96	0.72	-5.12	6.93
	HCQ9R	I doubt myself when I have a difficult task to undertake	3.19	1.09	-0.12	-2.09
	HCQ15R	I lack self-belief	3.44	1.09	-1.72	-1.51

Table 5.1. *(continued).*

Factor		Item	Mean	*S.D.*	S ratio	K ratio
Confidence - self-worth	HCQ5R	I doubt whether I am a worthwhile person	3.92	0.97	-2.65	-1.12
	HCQ17	I feel that I am a person of worth	3.92	0.84	-3.59	2.14
	HCQ21R	I take a negative attitude toward myself	3.63	1.11	-1.59	-2.26
	HCQ25	I feel that I am a truly worthwhile person	3.82	0.92	-3.28	1.31

Note: S ratio = Skewness ratio; K ratio = Kurtosis ratio.

Table 5.2. Correlations between factors of the HCQ and SDS and Cronbach's Alpha reliability coefficients using the exploratory sample.

	MT factors					
	CHAL	COM	CONT	CC	CSW	MTMT
CHAL	(0.60)					
COM	0.41**	(0.66)				
CONT	0.37**	0.39**	(0.59)			
CC	0.40**	0.54**	0.49**	(0.65)		
CSW	0.38**	0.39**	0.51**	0.64**	(0.82)	
MTMT	0.67**	0.74**	0.74**	0.80**	0.76**	(0.88)
SDS	0.13	0.13	0.12	0.01	- 0.03	0.11

Note: **Significance is at $p < 0.01$. Cronbach's Alpha reliability coefficients are displayed in parentheses. CHAL = Challenge; COM = Commitment; CONT = Control; CC = Confidence - competence; CSW = Confidence – self-worth; MTMT = Mean Total MT; SDS = Social Desirability Scale.

Table 5.3. EFA factor loadings of the secondary four factor model of the HCQ.

Identified factor	Associated variables	RFL
Confidence	HCQ1	0.740
	HCQ5R	0.597
	HCQ15R	0.651
	HCQ17	0.783
	HCQ21R	0.643
	HCQ25	0.784
Commitment	HCQ6R	0.714
	HCQ8R	0.725
	HCQ11R	0.767
	HCQ23R	0.723
Control	HCQ4	0.596
	HCQ16	0.546
	HCQ19R	0.638
	HCQ22	0.636
Challenge	HCQ3R	0.433
	HCQ10	0.768
	HCQ13	0.695
	HCQ20	0.450

The HCQ items retained in the first factor were defined as confidence and accounted for 25.8% of the variance. Factor loadings dictated that HCQ1, HCQ5R, HCQ15R, HCQ17, HCQ21R, and HCQ25 should be grouped together. The second component, commitment, accounted for 9.8% of the variance. Commitment comprised of items HCQ6R, HCQ8R, HCQ11R, and HCQ23R. The third component, control, was made up of HCQ4, HCQ16, HCQ19R, and HCQ22 and accounted for 8.1% of the variance. The final component, challenge, was made up of HCQ3R, HCQ10, HCQ13, and HCQ20 and accounted for 6.1% of the variance. In addition to those retained items, a number of items were found to cross- and mis-load with an unhypothesised factor and were subsequently removed. For challenge, HCQ26 cross-loaded with the control component and HCQ24R was found to mis-load on to the confidence component. For commitment, HCQ2 was found to mis-load on to the challenge component and HCQ18 was found to mis-load on to the control component. For control, HCQ27R cross-loaded with the commitment component, HCQ12 was found to mis-load on to the confidence component and HCQ14R mis-loaded on to the commitment component. For confidence – competence, HCQ7 mis-loaded on to the control component and HCQ9R was found to mis-load onto the commitment component. No items for the confidence – self-worth component were found to cross- or mis-load (see Appendix 3.16 for the complete HCQ rotated factor loadings).

The findings from the respective rotated factor loadings therefore identified a four factor model of MT whereby the confidence – competence and confidence – self-worth components were extracted as a single component, named confidence. Given that the content of the confidence – competence and confidence - self-worth constructs were captured (2 items and 4 items, respectively), it was decided to accept the extracted four factor model of MT. Factor loadings yielded by the EFA revealed that in general, items relating to the commitment component were found to have the strongest relationships with their associated component and that items relating to the challenge component were found to have the weakest relationships with their associated component (see Table 5.3). Inspection of factor loadings shows that 15 out of 18 items (83%) could be considered to have good or above relationships with their associated components (see Table 5.3). Specifically, factor loadings obtained from the individual items of the confidence and commitment components were generally considered to have very good or above relationships with their associated components, with 5 out of 6 (83%) and 4 out of 4 (100%) considered very good or above, respectively. Factor loadings obtained from the individual items of the control component were considered to have good to very good relationships with their associated component, with 2 out of 4 (50%) considered good and the remaining items considered very good. Factor loadings obtained from the individual items of the challenge component were diverse given that 2 out of 4 items (50%) were considered to have very good to excellent relationships with their associated component, yet the remaining items could only be considered to be fair to good (see Appendix 3.16). This resulted in an 18-item four factor model of MT being taken forward to be analysed by CFA. This model will be known as the revised 4C's model of MT hereafter.

Confirmatory sample

Descriptive statistics

Descriptive statistics of HCQ scores using the confirmatory sample are depicted in Table 5.4. Analysis of skewness and kurtosis ratios revealed that the data were generally nonnormally distributed. Inter-factor correlations of the HCQ were weak to moderate suggesting that the factors represent related yet independent components of the revised 4C's model of MT (see Table 5.5). The correlations between the respective factors and the SDS were generally weak which suggests that socially desirable responding had little impact on the regression weights observed (see Table 5.5). The HCQ demonstrated adequate internal reliability in two out of the four factors measured ($\alpha > 0.70$: Nunnally & Bernstein, 1994; commitment: 0.80; confidence: 0.85). Overall, the MTQ48 demonstrated high internal reliability ($\alpha = 0.86$). Table 5.5 displays the respective Cronbach's Alpha coefficients.

Confirmatory Factor Analyses

Goodness of fit

The CFA's conducted on the hypothesised first- and second-order revised four factor models of MT revealed inadequate model fit relative to the conservative and liberal guidelines. Table 5.6 indicates that although the first-order revised model provided marginally better model fit than the second-order revised model, in general, the respective fit indices are below the conservative thresholds (identified in parentheses).

Regression weights

Regression weights of the first- and second-order hypothesised four factor model of MT revealed that in general, items relating to the commitment factor were found to have the strongest relationships with their hypothesised factor and that items relating to the challenge factor were found to have the weakest relationships with their hypothesised factor. Inspection of factor estimates showed that 12 out of 18 items (67%) could be considered to have good or above relationships with their hypothesised factors. Specifically, estimates obtained from the commitment and confidence factors were considered to have very good to excellent relationships with their hypothesised factors, with 4 out of 4 items (100%) and 5 out of 6 items (83%) considered very good or above, respectively. Estimates obtained from the individual items of the control factor were diverse in that 2 out of 4 items (50%) were considered to have very good or above relationships with their hypothesised factor, yet the remaining items were considered to have fair or below relationships. Estimates obtained from the challenge factor were considered to have fair or below relationships with their hypothesised factor, with 3 out of 4 items (75%)

considered to have fair or below relationships. However, it is important to note that the remaining challenge item was considered to have an excellent relationship with its hypothesised factor. All estimates were statistically significant at the $p < 0.05$ level (see Appendix 3.17).

Model re-specification

Due to the insufficient model fit of the respective hypothesised models of MT, the item with the weakest relationship with its hypothesised factor was removed in an effort to improve model fit. This resulted in the removal of item HCQ3R (challenge), item HCQ6R (commitment), item HCQ22 (control), and item HCQ1 (confidence). This resulted in the revised models being represented by 14 items (challenge: $n = 3$; commitment: $n = 3$; control: $n = 3$; confidence: $n = 5$). Table 5.6 summarises the fit indices of the originally hypothesised models and the revised (re-specified) models. Findings showed that although the first- and second-order four factor revised models showed the same level of change in CFI (ΔCFI) from its respective hypothesised model specification, the first-order revised model provided the best fitting model. Despite the respective fit indices moving towards the conservative thresholds, and in some cases reaching the thresholds, the overall improvements in model fit did not provide an adequately fitting model in accordance with the most conservative thresholds (see Study 1 Method).

However, when the criteria for assessing model fit were offset to more liberal thresholds, the CMIN/DF (< 3.00; Ullman, 2001), RMSEA (< 0.08; Browne & Cudeck, 1993), and CFI (> 0.90; Bentler, 1992) could be considered to represent adequate model fit. Additionally, the PCFI could be considered to represent good model fit when using the most conservative threshold (close to or above 0.60; Blunch, 2008). Given that the AIC decreased substantially from the hypothesised first-order four factor model (4448.494) to the revised first-order four factor model (254.395), the fit indices of the revised first-order four factor model appear to represent acceptable model fit. Table 5.7 displays the items making up the revised first-order four factor model of the HCQ and their respective regression weights. Inspection of regression weights revealed that items were generally strong indicators in that 11 out of 14 items could be considered to have very good to excellent relationships with their hypothesised factors. Significant and weak to moderate inter-factor correlations provided support for item convergent and discriminant validity (factors were related yet distinct), which in turn, provided support for its acceptance (see Appendix 3.18 for inter-factor correlations).

Table 5.4. Descriptive statistics of the HCQ by factor using the confirmatory sample.

Factors		Item	Mean	*S.D.*	S ratio	K ratio
Challenge	HCQ3R	I do not look forward to challenging situations	3.56	0.98	-5.33	1.57
	HCQ10	I believe that change enables me to grow as a person	3.97	0.77	-5.36	5.15
	HCQ13	I see challenges in my life as opportunities for me to develop as a person	4.05	0.67	-2.69	1.78
	HCQ20	Changes in my daily routine encourage me to learn	3.35	0.74	-3.15	1.26
Commitment	HCQ6R	I often find myself disengaging from tasks	3.16	0.96	-1.13	-2.03
	HCQ8R	I find it difficult to stay committed to whatever I am doing	3.76	0.88	-5.06	2.67
	HCQ11R	I find it difficult to stay committed to achieving my goals	3.78	0.88	-5.92	3.22
	HCQ23R	I lack commitment	4.00	0.75	-4.46	2.39
Control	HCQ4	I can influence the path that my life takes	4.14	0.64	-3.42	5.81
	HCQ16	Events in my life are shaped by my own actions	3.89	0.74	-4.02	4.36
	HCQ19R	My actions have little influence on my life	3.95	0.84	-6.27	4.30
	HCQ22	What happens to me is my own doing	3.79	0.77	-3.96	2.25
Confidence	HCQ1	I am confident in my own abilities	3.98	0.74	-5.41	4.03
	HCQ15R	I lack self-belief	3.67	1.07	-4.82	0.04
	HCQ5R	I doubt whether I am a worthwhile person	4.02	0.94	-4.69	0.38
	HCQ17	I feel that I am a person of worth	3.85	0.75	-4.31	3.20
	HCQ21R	I take a negative attitude toward myself	3.76	1.00	-5.01	0.92
	HCQ25	I feel that I am a truly worthwhile person	3.85	0.81	-6.24	6.17

Note: S ratio = Skewness ratio; K ratio = Kurtosis ratio

Table 5.5. Correlations between factors of the HCQ and SDS scores and Cronbach's Alpha reliability coefficients using the confirmatory sample.

	MT factors				
	Challenge	Commitment	Control	Confidence	MTMT
Challenge	(0.48)				
Commitment	0.46**	(0.80)			
Control	0.32**	0.20**	(0.63)		
Confidence	0.41**	0.43**	0.41**	(0.85)	
MTMT	0.68**	0.67**	0.60**	0.85**	(0.86)
SDS	0.27**	0.35**	0.11	0.08	0.23**

Note: **Significance is at $p < 0.01$. Cronbach's Alpha Reliability Coefficients are displayed in parentheses.
MTMT = Mean Total MT; SDS = Social Desirability Scale.

Table 5.6. Summary of fit indices across hypothesised model specification, model re-specification and structural models.

	Fit indices						
CFA model	CMIN/DF	*P*	CFI	Δ CFI	PCFI	RMSEA (90% CI)	AIC
Criterion values	(< 2.00)	(> 0.05)	(> 0.95)		(> 0.60)	(< 0.05)	(lower = better)
Hypothesised models							
First-order four factor	2.546	0.000	0.835	-	0.704*	0.086 (0.075, 0.098)	4448.494
Second-order four factor	2.616	0.000	0.825	-	0.707*	0.088 (0.077, 0.100)	458.670
Revised models							
First-order four factor	2.231	0.000	0.904	0.069	0.705*	0.077 (0.061, 0.093)	254.395*
Second-order four factor	2.320	0.000	0.894	0.069	0.717*	0.080 (0.064, 0.095)	261.346*
Structural models							
HPC	2.373	0.000	0.892	-0.012[1]	0.706*	0.081 (0.066, 0.097)	264.829
CPH	3.146	0.000	0.824	-0.080[1]	0.679*	0.102 (0.087, 0.116)	323.917

Note: *denotes good fit; [1] Δ CFI for HPC and CPH models are calculated using CFI from the first-order four factor revised model.

Structural models

Goodness of fit

The HPC and CPH structural models of the first-order four factor revised model of the HCQ revealed inadequate model fit when using the more conservative and more liberal model fit thresholds. Table 5.6 indicates that although the HPC structural model of the first-order four factor revised model of the HCQ provided marginally better model fit than the CPH, in general, the respective fit indices are below the required thresholds (identified in parentheses). Overall, the fit indices of the structural models revealed inferior model fit when compared to the heterarchical first-order four factor revised model of the HCQ.

Regression weights

Inspection of regression weights of the HPC and CPH structural models shows that 11 out of 14 (79%) and 12 out of 14 (86%) items could be considered to have good or above relationships with their hypothesised factors, respectively. All estimates were statistically significant ($p <$ 0.05). Regression weights for the HPC and CPH structural models are displayed in Figure 5.1 and Figure 5.2.

Table 5.7. Standardised regression weights of the first-order four factor revised model of the HCQ.

Item / factor			Estimate
HCQ10	<---	Challenge	0.864
HCQ13	<---	Challenge	0.467
HCQ20	<---	Challenge	0.385
HCQ11R	<---	Commitment	0.763
HCQ23R	<---	Commitment	0.762
HCQ8R	<---	Commitment	0.718
HCQ4	<---	Control	0.702
HCQ16	<---	Control	0.676
HCQ19R	<---	Control	0.474
HCQ25	<---	Confidence	0.800
HCQ21R	<---	Confidence	0.743
HCQ5R	<---	Confidence	0.719
HCQ17	<---	Confidence	0.708
HCQ15R	<---	Confidence	0.699

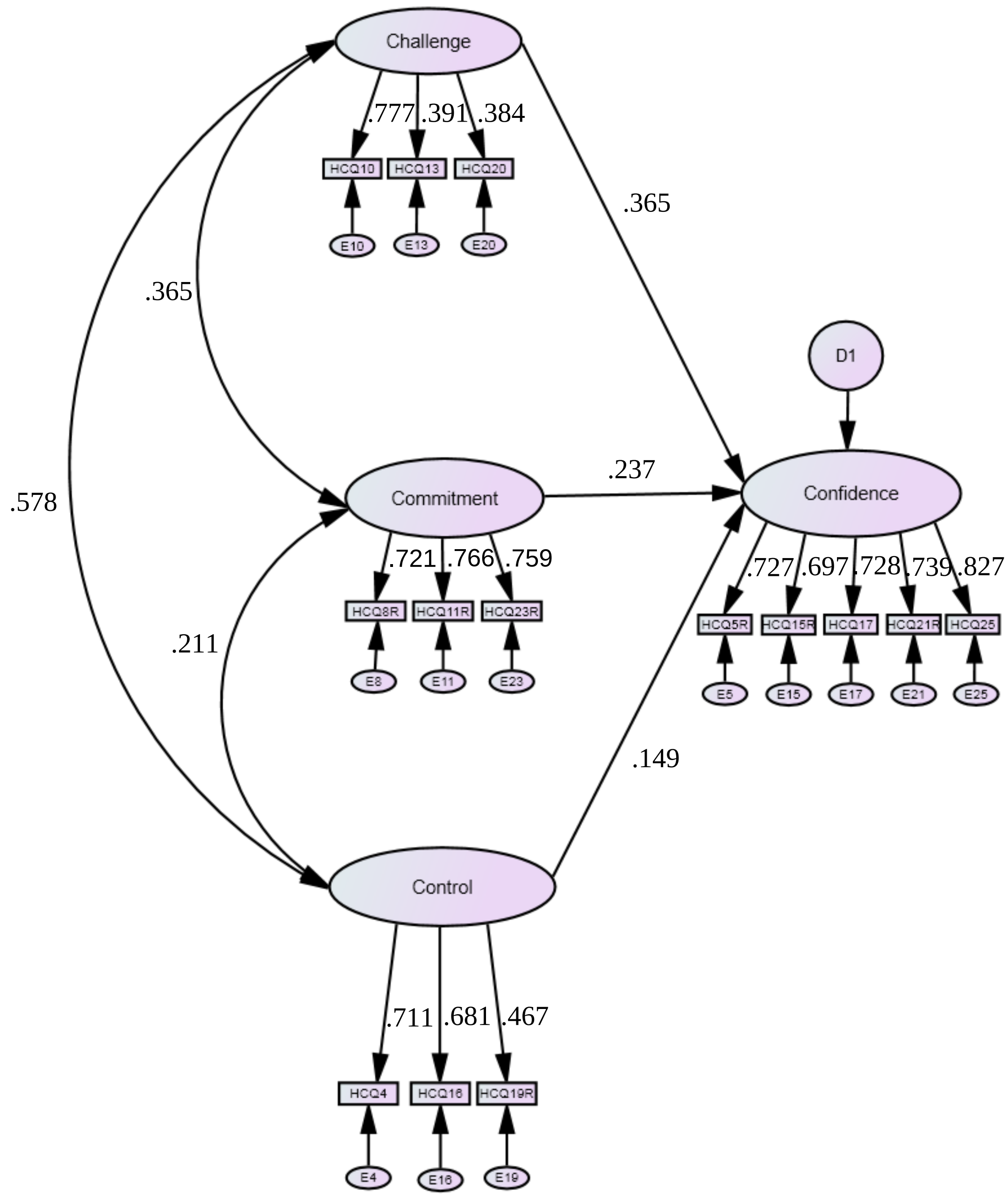


Figure 5.1. Path diagram depicting the HPC structural model of the HCQ.

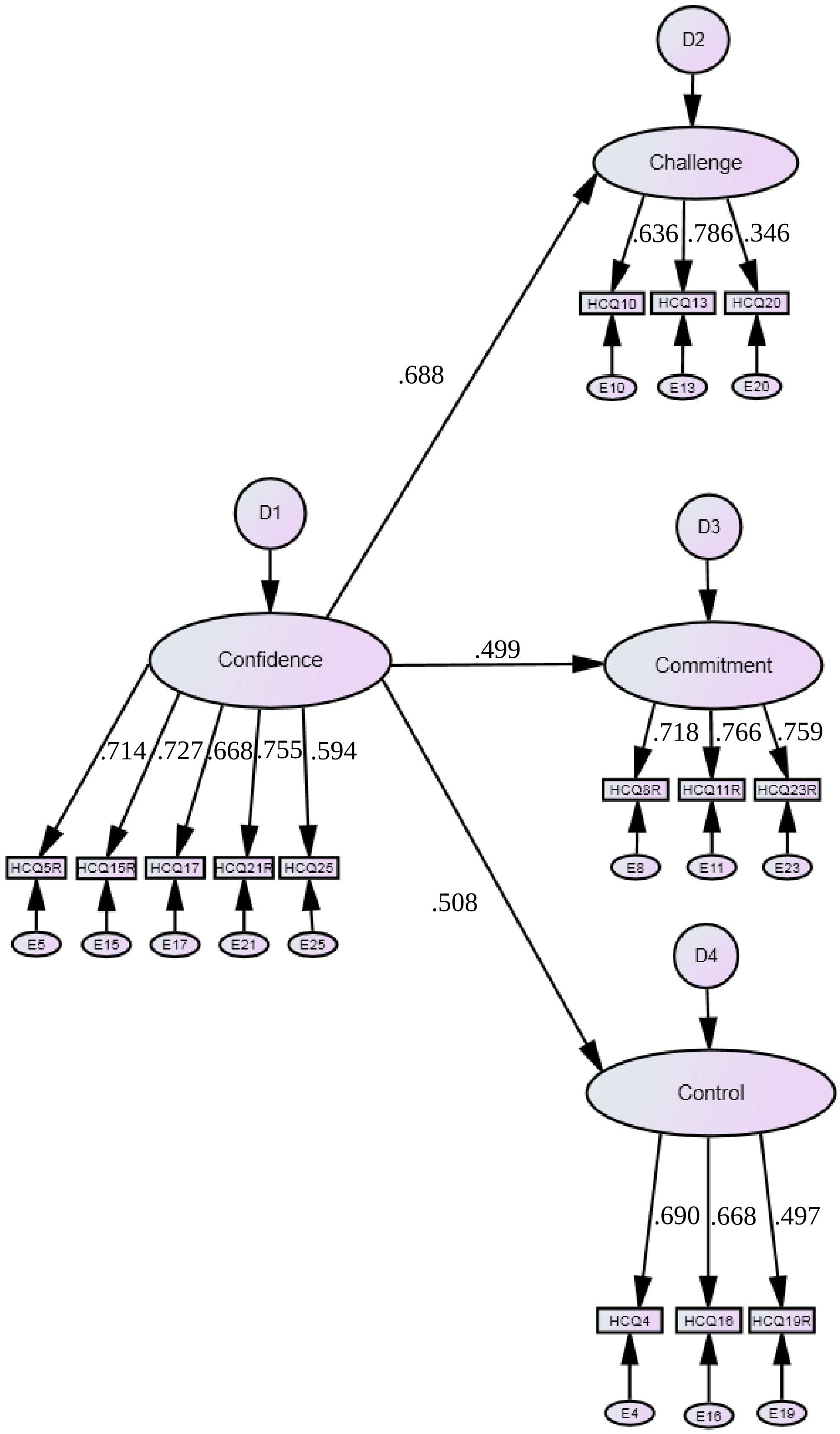


Figure 5.2. Path diagram depicting the CPH structural model of the HCQ.

5.4 Discussion

The first aim of Study 3 was to regenerate definitions which better represent the focal constructs underpinning Clough *et al.'s* (2002) 4C's model of MT. Inspection of the hardiness literature provided definitions which adequately represented the challenge, commitment, and control constructs. Given the limited research which has examined dispositional confidence and the lack of theoretical underpinning for this construct (Bandura, 1997), the use of self-esteem theory to construct the confidence definition appears to have provided a sound conceptual underpinning to its development. In line with the recommendations outlined in the scale development literature (e.g., Clark & Watson, 1995; MacKenzie *et al.*, 2011), the definitions constructed in Study 3 appear to provide adequate representations of their focal domains in regards to their content and structure. The revised model of MT therefore appears to provide a conceptually embedded model to examine the traits thought to underpin MT.

The second aim of Study 3 was to generate items which accurately represent the factor definitions of the revised 5C's model of MT. Inspection of item CVI's revealed that 25 out of the initial 39 items (63%) generated exceeded Lynn's (1986) recommended cut-off (close to 0.80 when using six scrutineers) for item acceptance. In addition, 16 out of the 25 items (64%) accepted were found to have CVI's of 1 indicating that all six expert scrutineers rated the item as an excellent match to its hypothesised factor definition. Analysis of item CVI's therefore provides support for the generated items to accurately represent the revised 5C's model of MT. Consequently, the methods employed in Study 3 appear to satisfy the objectives of the construct and item generation processes outlined in the scale development literature (e.g., Clark & Watson, 1995; MacKenzie *et al.,* 2011).

The third aim of Study 3 was to examine the factorial validity of the HCQ. The findings of the secondary EFA identified that four components should be extracted (confidence, commitment, control, challenge) and taken forward for further analysis. This finding was therefore inconsistent to the make-up of the revised 5C's model of MT whereby self-confidence was hypothesised to be made-up of two nested components; confidence – competence and confidence – self-worth. It was decided to accept the extracted four factor model of MT (revised 4C's model of MT) on the grounds that the retained items represented both confidence - competence and confidence self - worth. This finding shows that although the extracted four factor model represented the confidence construct as a single component, its make-up was actually capturing the conceptual content which it intended to capture (Clark & Watson, 1995; MacKenzie *et al.*, 2011). Moreover, the extracted four factor model was supported by the extent to which the confidence component explained the model in that the secondary EFA showed the confidence component to account for 25.8% of the variance, which was superior to all other components extracted. The rotated factor loadings observed for the secondary EFA revealed that the relationships between the items and the challenge, commitment, control, and confidence components were relatively strong in that 2 out of 4 (50%), 4 out of 4 (100%), 2 out of 4 (50%), and 5 out of 6 (83%) items could be considered to have very good or above relationships with their associated components. The findings from the secondary EFA therefore indicated that the extracted items were forming coherent subsets in the data which were relatively independent of one another (Tabachnick & Fidell, 2001). This finding was supported by the significant weak to moderate inter-factor correlations of the HCQ (see Figure 5.2), which in turn provides support for the discriminant validity between its respective factors.

Despite the secondary EFA indicating support for the revised 4C's model of MT, CFA's of the hypothesised first- and second-order models of the HCQ revealed inadequate fit when using the most conservative thresholds (see Table 5.6). However, findings provided preliminary support for the studies hypotheses in that model re-specifications revealed that when using more liberal thresholds, fit indices of the first-order four factor revised model did show acceptable fit to the data (see Table 5.6). The use of goodness of fit indices has received ample debate in the scale validation literature, with some researchers (e.g., Hu & Bentler, 1999; Russell, 2002) emphasising the need to use more stringent thresholds, whereas other researchers (e.g., Marsh *et al.,* 2004) have questioned this approach and have suggested using more liberal thresholds to evaluate model fit. Indeed, previous research examining hardiness has used more liberal model fit thresholds to support factorial validity. Specifically, when examining the factorial validity of the DRS15-R, Hystad *et al.* (2010) used liberal thresholds (CFI, Goodness of Fit Index [GFI], Incremental Fit Index [IFI] $\geq$ 0.90 and SRMR $\leq$ 0.08) to accept the second-order three factor model of hardiness (CFI = 0.92, GFI = 0.93, IFI = 0.92, SRMR = 0.068). The findings of Study 3 are therefore comparable to data used to support the factorial validity of existing hardiness measures. Furthermore, researchers (e.g., Byrne, 2010) have stated that goodness of fit indices are only one component of the model evaluation process and have emphasised the importance of using regression weights to assess model adequacy. Given that 11 out of 14 HCQ items could be considered to have very good to excellent relationships with their hypothesised factors, the arguments of Byrne and Marsh *et al.* could be used to indicate preliminary support for the factorial validity of the first-order four factor revised model of the HCQ.

Given the importance placed on examining the dimensionality of an instrument's conceptual model (e.g., Gignac, 2009; Marsh *et al.,* 2010), Study 3 also examined alternative models to underpin the HCQ. The findings offered little support for the factorial validity of the HPC and CPH structural models in that model fit was inadequate when using the most liberal and most conservative model fit thresholds (see Table 5.6). The lack of support for the factorial validity of the structural models suggests that the constructs underpinning the HCQ are best represented by the heterarchical first-order four factor revised model as opposed to the HPC and CPH hierarchical models. This finding indicates that the constructs underpinning the revised 4C's model of MT are related, yet independent factors which contribute equally to the traits thought to underpin MT.

The first-order dimensionality of the four factor revised model of the HCQ is therefore inconsistent with the theoretical propositions of hardiness (Kobasa, 1979) and previous inventories measuring hardiness (Hystad *et al.*, 2010; Maddi *et al.*, 1999) whereby the subcomponents of hardiness are accounted for by a higher-order hardiness construct. The first-order dimensionality of the HCQ therefore supports the contentions of Carver (1989) whereby using scores to represent the subcomponents of hardiness may be a more appropriate means of interpreting and understanding hardiness scores as opposed to assessing combined scores. Future research could use HCQ subscale scores to examine whether the separate traits have different effects on cognitive and behavioural outcomes. A particularly fruitful avenue for further research could be to examine the independent and interactive impact of the traits thought to underpin MT by ascertaining their ability to predict cognition and behaviour in sport.

In addition to factorial validity and dimensionality, the distributional properties of an instrument have important implications for the evaluation of within-network validity (Marsh, 2002). The relatively high item mean scores obtained in Study 3 resulted in the identification of negatively skewed data, showing that participants generally reported high levels of dispositional challenge, commitment, control, and confidence. Previous research (e.g., Cumming, Clark, Ste-Marie, McCullagh, & Hall, 2005; Hall, Mack, Paivio, & Hausenblas, 1998; Williams & Cumming, 2011) has used response variability as a means to assess an instrument's distributional properties. Specifically, item standard deviations greater than 1 are deemed satisfactory. Despite the data in Study 3 being negatively skewed, item standard deviations approached or were greater than 1, thus providing preliminary support for the HCQ-R's distributional properties. It is important to note, however, that high mean scores (negative skewness) should not be used to draw inferences regarding an instrument's factorial validity whereby assessment should primarily focus on model fit and factor estimates (Byrne, 2010).

A possible limitation is the small number of items included in the HCQ. Although smaller item questionnaires aid practical use and facilitate the independence of factors (Burisch, 1997), it raises the possibility that the instrument may be limited in its ability to adequately capture the traits which it intends to capture. However, given that each factor of the HCQ is represented by at least three items, one may argue that the content validity of the instrument is not sufficiently compromised to cause concern. Despite there being a preference for 4-item factors for questionnaire development (Clark & Watson, 1995; Kline, 2000; MacKenzie *et al.*, 2011), previous instruments containing 3-item factors that have been rigorously examined are present in the literature (e.g., Achieve Goals Questionnaire for Sport; Conroy, Elliot, & Hofer, 2003). Although research has emphasised the need to examine factorial invariance once the factor structure has been established (e.g., Cheung & Rensvold, 2002; Horn & McArdle, 1992; Marsh, 2002), the representation (i.e., ratio of males vs. females) and size of the sample used in Study 3 was deemed inadequate to afford such analyses. Research by Meade (2005) examined the impact of sample size on invariance testing and found samples of 100 and 200 to be inadequately powered when compared to a sample of 400. The cross-sectional design of Study 3 prevented the assessment of test-retest reliability, which has been argued to be an important consideration when examining the construct validity of personality constructs (Marsh; McCrae *et al.*, 2011). Future research is therefore required to examine the test-retest reliability of the HCQ to further enhance our understanding of its within-network properties.

Using the scale development literature as a guiding framework (e.g., MacKenzie *et al.*, 2011), Study 3 provides the development and psychometric evaluation of an instrument designed to measure the traits thought to underpin MT. Although the internal consistency estimates for HCQ challenge and control (confirmatory sample) were below the required threshold ($\alpha > 0.70$; Nunnally & Bernstein, 1994), Study 3 presents preliminary support for the factorial validity and dimensionality of the HCQ when using more liberal model fit thresholds. One may therefore argue that Study 3 has provided the necessary prerequisite to the examination of the HCQ's between-network properties (Gignac, 2009; Marsh & Hau, 2007). Consistent with the construct validation approach (Marsh, 2002), research is warranted to examine the HCQ's between-network properties. Research examining convergent and predictive validity would enhance our understanding of the HCQ's ability to measure the traits which it intends to capture and the importance of the separate traits in predicting cognitions and behaviours in sport.

5.5 Summary

The aim of Study 3 was to regenerate definitions and items which accurately represent the core traits underpinning Clough *et al.'s* (2002) 4C's model of MT in an effort to develop a valid measure of MT. Although inspection of the regenerated factor definitions provided support for their representation of the traits thought to make-up MT and that analysis of item content provided support for their representation, the findings revealed little support for the factorial validity of the hypothesised models of the HCQ when using the most conservative model fit thresholds. However, preliminary support was found for the first-order four factor revised model of the HCQ in that fit indices satisfied more liberal thresholds. Given that researchers have supported the use of more liberal model fit thresholds (Marsh *et al.*, 2004) and the use of both fit indices and factor loadings when assessing factorial validity (Byrne, 2010), the findings of Study

3 provide preliminary evidence to support the HCQ's factorial validity and dimensionality. Study 4 will aim to extend the psychometric evaluation of the HCQ by examining its test-retest reliability, convergent validity, and predictive validity.

CHAPTER 6.0

STUDY 4: A PRELIMINARY EXAMINATION OF THE CONSTRUCT VALIDITY OF THE REVISED HARDINESS CONFIDENCE QUESTIONNAIRE

6.1 Introduction

The findings from Study 3 provided preliminary evidence to support the factorial validity of the first-order four factor revised model of the HCQ (referred to as the HCQ-R hereafter) when using liberal model fit thresholds. In an effort to further examine the HCQ-R's factorial validity, Study 3 used structural models to examine the relations among its underlying latent constructs yet found superior model fit for the heterarchical dimensionality of the HCQ. Given that preliminary factorial support was found for the HCQ-R, Study 4 sought to further examine the construct validity of the measure. According to Marsh (2002), construct validation requires the evaluation of both within- and between network properties. The aim of Study 4, therefore, was to further examine the within- and between-network properties of the HCQ-R by examining its test-retest reliability, convergent validity, and predictive validity.

Test-retest reliability

Researchers (e.g., MacKenzie *et al.,* 2011; Marsh, 2002) have emphasised the importance of assessing reliability when evaluating an instruments psychometric properties. An important aspect of reliability is the extent to which constructs remain stable over time (Lane *et al.,* 2005). Test-retest reliability is considered to be an effective means of assessing the stability of test scores (Anastasi & Urbina, 1997; Bland & Altman, 1986; Kline, 1993; Nevill, 1996; Nevill, Lane, Kilgour, Bowes, & Whyte, 2001; Wilson & Batterham, 2001). Research in mainstream psychology has generally supported the contention that personality traits are relatively stable and evolve over time (Mathews & Deary, 1999; McCrae *et al.,* 2000). Test-retest reliability is therefore considered to have particular importance when evaluating instruments which assess personality traits (McCrae *et al.,* 2011; Watson, 2004). Intraclass correlations coefficients have been widely used in the sport and exercise psychology literature (e.g., Lonsdale *et al.,* 2008; Williams & Cumming, 2011) to assess an instrument's test-retest reliability. Despite Anastasi and Urbina and Kline stating that test-retest correlations greater than 0.80 demonstrate acceptable levels of stability, Vincent (1999) has suggested that correlations greater than 0.70 provide acceptable levels of stability.

Convergent validity

Many researchers consider assessing the convergent validity of an instrument to be vital in the construct validation process (e.g., MacKenzie *et al.,* 2011; Marsh, 2007). The assessment of convergent validity involves using instruments that purport to measure the same or substantially overlapping constructs with the instrument being examined, with strong or substantial relationships ($r > 0.07$) indicative of convergent validity (MacKenzie *et al.*, 2011; Marsh, 2007). Given that the HCQ-R is underpinned by the revised 4C's model of MT, which in turn is made-up of hardiness and confidence (self-esteem), Study 4 will use a measure of hardiness and a measure of self-esteem to examine the convergent validity of the HCQ-R.

Predictive validity

Predictive validity is an extension of convergent validity and is concerned with the extent to which scores from one measure predict scores from another measure which it is hypothesised to be related (Vaughn & Daniel, 2012). Constructs that are theoretically related to the focal construct should be obtained to establish an instrument's nomological network (e.g., Cronbach & Meehl, 1955). Evaluation of an instrument's nomological network is important as it enables one to a) understand the lawful relationships between the focal construct and other constructs, and b) test whether the indicators of the focal construct relate to measures of other constructs in the manner expected (Cronbach & Meehl; MacKenzie *et al.,* 2011; Marsh, 2002).

Inspection of the core tenets underpinning the HCQ-R indicates that it has theoretical associations with coping styles and stress appraisal. The theoretical predictions of hardiness argue that high levels of challenge, commitment, and control motivate one to respond to stressors with specific coping efforts which facilitate resiliency by turning potential disasters into opportunities (e.g., Kobassa, 1979; Maddi, 2002; Maddi & Kobassa, 1984). Maddi and Kobasa posit that strong hardy attitudes result in higher levels of courage and motivation, which in turn, facilitate carrying out problem-solving coping and building social support as means of effective self-care. More specifically, individuals who believe change can be used to stimulate personal growth (challenge), who are deeply involved in the activities they undertake (commitment), and who believe they can continually influence their environment (control) are more likely to be motivated to respond to stressful events by elevating their engagement with them (Kobassa; Maddi & Kobasa). Additionally, hardy individuals are thought to see stressful events through a lens which enables them to cognitively restructure the event so that it appears less threatening, which in turn, enables them to behave in a way where they can benefit and learn from the experience (Kobassa; Maddi & Kobasa). In contrast, individuals low in hardiness are thought to adopt more regressive coping because they have a tendency to feel threatened (rather than challenged) by change in one's environment, alienated from (rather than committed to) activities in which they undertake, and powerless (rather than in control) in face of the ever-changing environment in which they live (Kobassa; Maddi & Kobasa).

Research examining the theoretical predictions of hardiness has provided strong support for its positive relations with problem- and emotion-focused coping, and its negative associations with avoidance-focused coping. For instance, research (Goss, 1994; Maddi & Hightower, 1999;

Maddi, 1999) has found high levels of hardiness to be associated with problem-solving coping (e.g., active coping, seeking instrumental social support) and negatively associated with avoidance coping (e.g., mental disengagement, denial). Noting that information may be lost when only using total hardiness scores (Carver, 1989), researchers have examined the relationships between the respective hardiness subcomponents and coping styles. In general, commitment and control have shown moderate to strong positive correlations with problem- and emotion-focused coping and moderate to strong negative correlations with avoidance-focused coping (e.g., Florian, Mikulincer, & Taubman, 1995; Klag & Bradley, 2004; Wadey *et al.,* 2012; Williams, Wiebe, & Smith, 1992). Although associations between challenge and coping styles have been established, the strength of correlations obtained have generally been weak-to-moderate.

Despite researchers (e.g., Kobassa, 1979; Maddi & Kobassa, 1984) positing that hardiness motivates one to appraise potentially stressful situations as less threatening, research by Weibe (1991) is the only study to date which has explicitly examined this contention. Weibe found that high hardy individuals rated the same objective stressor as less threatening than low hardy individuals, with high hardy individuals reporting higher levels of control than their low hardy counterparts. This finding is consistent with theoretical models in the sport and exercise psychology literature whereby control is thought to determine challenge and threat states. Jones, *et al.'s* (2009) Theory of Challenge and Threat States in Athletes posits that perceived control regarding oneself and the environment determines resource appraisals and subsequent challenge and threat states in competition, with high levels of control indicative of challenge states and low levels of control indicative of threat states. Closely linked to stress appraisals is Jones' (1995) Control Model of Debilitative and Facilitative Competitive State Anxiety. According to Jones, facilitative anxiety results from an athlete's perception of control over the environment and the self, sufficient coping beliefs, and sufficient self-efficacy regarding goal attainment. Jones posits that pre-competitive emotions can be broadly categorised to two ways; anxiety symptoms are perceived as helpful or unhelpful to performance (c.f. Hanton, Neil, & Mellalieu, 2008). Accordingly, one may argue that anxiety interpretation offers an additional means to assess the HCQ-R's predictive validity.

Upon evaluating the theoretical predictions of hardiness, it is evident that coping styles, stress appraisals, and pre-competitive anxiety symptoms offer constructs within the nomological network of the revised 4C's model of MT. Assessing the relationships between MT and coping styles therefore provides a means of assessing the extent to which subscales of the HCQ-R may predict problem-, emotion-, and avoidance-focused coping strategies in sport. Furthermore, assessing the relationships between MT and challenge and threat appraisals and pre-competitive state anxiety symptoms provides a means of examining the HCQ-R's predictive validity.

The aim of Study 4, therefore, was to examine the test-retest reliability, convergent validity, and predictive validity of the HCQ-R. Study 4 presents two investigations to examine the psychometric properties of the HCQ-R. Investigation 1 details an evaluation of the HCQ-R's test-retest reliability, convergent validity, and predictive validity. Investigation 2 details an additional examination of the HCQ-R's predictive validity. Specifically, Investigation 1 examined the HCQ-R's predictive validity whereby scores were correlated with self-reported coping strategies used by athletes to deal with stressful situations in sport. A number of hypotheses were formulated. First, it was hypothesised that the test-retest reliability coefficients for the HCQ-R would be greater than 0.80 (Anastasi & Urbina, 1997; Kline, 1993) and thus demonstrate acceptable levels of stability. Second, it was hypothesised that HCQ-R subscale scores would correlate positively, significantly and strongly (r = 0.70; MacKenzie *et al.*, 2011; Marsh, 2007) with overlapping subscale scores of the DRS15-R (measure of hardiness) and the Revised Self-liking/Self-competence Scale (SLSC-R; measure of self-esteem) and thus demonstrate high convergent validity (MacKenzie *et al.*, 2012; Marsh, 2002). Third, it was hypothesised that HCQ-R scores would be positively and significantly correlated to problem- and emotion-focused coping strategies, and significantly and negatively correlated to the use of avoidance-focused strategies to deal with stressful events. It was also hypothesised that all four HCQ-R factors would significantly predict the use of coping strategies.

6.2 Investigation 1

6.2.1 Method

Participants

Green (1991) recommends a minimum sample of 50 + 8k (where k is the number of predictors) to conduct regression analyses. Given that there are four predictor variables (challenge, commitment, control, confidence) in Study 4, the minimum sample required was 82 participants. Consequently, one hundred competitive student athletes were recruited to participate in Investigation 1 of Study 4. Participants consisted of 72 males and 28 females (M age = 21.18 years, *S.D.* = 3.85) with a mean of 10.07 years (*S.D.* = 4.22) competitive playing experience in their primary sport. The vast majority of participants (95%) were White British, with the remaining participants (n = 5) consisting of a variety of ethnic backgrounds such as Black or Black British African (n = 2), Black or Black British Caribbean (n = 2), and Mixed White and Black African (n = 1). The primary sport of participation included both team sports (n = 72), such as Football (n = 32), Rugby (n = 14), Cricket (n = 11), and Hockey (n = 5), and individual sports (n = 28), such as Athletics (n = 5), Badminton (n = 5), Swimming (n = 4), and Tennis (n = 4). The highest level of primary sport participation ranged from Intervarsity (n = 8) through Club (n = 36), County (n = 35), Regional (n = 8), National (n = 6), and International (n =7) level.

Measures

Convergent validity

The Revised Hardiness Confidence Questionnaire

The HCQ-R is a 14-item revised instrument of the original 27-item HCQ designed to measure the revised 4C's model of MT (see Appendix 4.1 for the HCQ-R). The procedures adopted to construct the HCQ-R operational document were identical to those used in Study 2. The HCQ-R requires respondents to rate their agreement to statements on a 5-point Likert scale ranging from (1) strongly disagree to (5) strongly agree. Participants were asked to respond to each item by considering how they are generally. The 14-item HCQ-R measures four subscales of challenge (3 items), commitment (3 items), control (3 items), and confidence (5 items). HCQ-R scores were reversed accordingly (see Appendix 4.2 for HCQ-R scoring key). Four subscales scores were obtained by calculating the mean of all items comprising each of the challenge, commitment, control, and confidence subscales. The evidence yielded from Study 3 provides preliminary support for the factorial validity and dimensionality of the HCQ-R. Cronbach's Alpha reliability coefficients for the HCQ-R are displayed in Table 6.2.

The Revised Dispositional Resilience Scale

The DRS15-R (Bartone, 1995) is a 15-item revised instrument of the original 45-item Dispositional Resilience Scale (Bartone, Ursano, Wright, & Ingraham, 1989) designed to measure the 3C's of hardiness (challenge, commitment, control; see Appendix 4.3 for the DRS15-R). The DRS15-R requires respondents to rate their agreement to statements on a 4-point Likert scale ranging from (0) not at all true to (3) completely true. The DRS15-R measures three subscales of challenge (5 items), commitment (5 items), and control (5 items). Example items include "changes in my routine are interesting to me" (challenge); "life in general is boring for me" (commitment); "by working hard you can nearly always achieve your goals" (control). The DRS15-R contains six negatively worded items. Scores were reversed accordingly. Three subscales scores were obtained by summing the respective items for each of the challenge, commitment, and control subscales (see Appendix 4.4 for DRS15-R scoring key). Research by Hystad *et al.* (2010) has provided support for the factorial validity of the second-order model of the DRS15-R (CFI = 0.92; GFI = 0.93; IFI = 0.92, SRMR = 0.07) and internal reliabilities (except challenge) of the DRS15-R in a Norwegian Military sample.

The Revised Self-liking/Self-competence Scale

The SLSC-R (Tafarodi & Swann, 2001) is a 16-item revised instrument of the original 20-item Self-liking/Self-competence scale (Tafarodi & Swann, 1995; see Appendix 4.5 for the SLSC-R). The SLSC-R is designed to measure global self-esteem and requires respondents to rate their agreement to statements on a 5-point Likert scale ranging from (1) strongly disagree to (5) strongly agree. The SLSC-R measures two subscales of self-liking (8 items) and self-competence

(8 items). Example items include "I tend to devalue myself" (self-liking) and "I am highly effective at the things I do" (self-competence). The SLSC-R contains eight negatively worded items. Scores were reversed accordingly (see Appendix 4.6 for SLSC-R scoring key). Two subscale scores were obtained by summing the respective items of each of the self-liking and self-competence subscales. Research by Tafarodi and Swann (2001) has provided support for the factorial validity (CFI = 0.92; NNI = 0.91; NI = 0.91; RMSEA = 0.06) of the SLSC-R's two-factor model (self-liking, self-competency) and its convergent and discriminant validity. Tafarodi and Swann obtained adequate internal reliability estimates for males and females for the self-liking (both $\alpha = 0.90$) and self-competency ($\alpha = 0.83$ and 0.82, respectively) subscales.

Social Desirability Scale

See Study 1 for details of the SDS.

Predictive validity

The Coping Function Questionnaire

The Coping Function Questionnaire (CFQ; Kowalski & Crocker, 2001) is an 18-item instrument designed to measure coping function in adolescent sport participants (see Appendix 4.7). The CFQ requires respondents to select a stressful situation in sport in the previous 12 months and rate how much they used each coping function on a 5-point Likert scale ranging from (1) not at all used to (5) very much used. The CFQ measures three subscales of coping including problem-focused coping (6 items), emotion-focused coping (7 items), and avoidance coping (5 items). Example items include "I tried to find a way to change the situation" (problem-focused); "I tried to view the situation in a way that made it seem less stressful" (emotion-focused); I tried to get out of the situation to get away from the stress" (avoidance). Coping function scale scores are determined by taking the mean of all items comprising each scale, with higher scores reflecting greater coping (see Appendix 4.8 for CFQ scoring key). Research by Kowalski and Crocker has provided support for the factorial (boys/girls: TLI = 0.90/0.88; CFI = 0.91/0.90; Robust CFI [RCFI] = 0.92/0.91; RMSEA = 0.07/0.08), convergent, and divergent validity of the CFQ. Kowalski and Crocker obtained adequate internal reliability for the problem-focused and emotion-focused subscales ($\alpha = 0.79$ and 0.80, respectively) but not for avoidance ($\alpha = 0.61$). The CFQ has been used in adult samples (e.g., Allen, Greenlees, & Jones, 2011; Hanton, Neil, Mellalieu, & Fletcher, 2008).

Procedure

The procedure used for Study 4, Investigation 1 was identical to Study 1. Participants were recruited over a two week period. Participant information sheet, consent form, demographic questionnaire, and written debrief are presented in Appendix 4.9. Given the predictive inadequacies of using cross-sectional designs when assessing an instrument's predictive validity

(Gignac, 2009), Study 4 utilised two data collection points whereby participants completed the HCQ-R and SDS one week prior to completing the DRS15-R, SLSC-R, and CFQ. Participants also completed the HCQ-R at the second time point to afford the calculation of test-retest reliability.

Data analysis

Data screening

Analysis of skewness and kurtosis ratios for HCQ-R scores across sitting one and sitting two revealed that the data were marginally nonnormally distributed (see Table 6.1). Specifically, HCQ-R scores across sitting one and sitting two are negatively skewed showing that the data are distributed towards the upper end of the HCQ-R scale (as depicted by the item means). Inspection of skewness and kurtosis ratios for the DRS15-R, SLSC-R, and the CFQ revealed that the measures were generally normally distributed (see Appendix 4.10).

Test-retest validity

Intraclass correlation coefficients with absolute agreement were calculated using a two-way random effect model (Ntoumanis, 2001) to assess test-retest reliability. Intraclass correlations were used as opposed to correlations because they take into account mean score changes across occasions and correct for chance agreement (Shrout & Fleiss, 1979). In light of the data being nonnormally distributed, ANOVA with Friedman Chi-square were calculated to examine differences among the mean scores of the HCQ-R across sitting one and sitting two. In line with the recommendations of Nevill *et al.* (2001) and Wilson and Batterham (2001), intraclass correlation coefficients were conducted for individual items of the HCQ-R.

Convergent and predictive validity

Correlations were conducted between the overlapping subscales of the HCQ-R, DRS15-R, and SLSC-R to examine the HCQ-R's convergent validity. Correlations were also conducted between the HCQ-R and the CFQ to assess predictive validity. Given that minimal associations were found between the HCQ-R subscale scores and CFQ subscale scores, regression analyses and the relevant data screening procedures were not performed. Data screening was conducted using SPSS statistics 20.0.

6.2.2 Results

Descriptive statistics

Descriptive statistics of HCQ-R item scores across sitting one and sitting two are depicted in Table 6.1. Descriptive statistics by factor of the HCQ-R (sitting one and sitting two), SDS, DRS15-R, SLSC-R, and CFQ are displayed in Table 6.2. The correlations between the respective factors of the HCQ-R and the SDS were generally weak suggesting that socially desirable responding had little impact on the scores obtained for the HCQ-R. Although only 2 out of the 4 factors measured by the HCQ-R (across both sitting one and two) demonstrated adequate internal reliability ($\alpha > 0.70$; Nunnally & Bernstein, 1994), the estimates obtained in sitting two did approach the required threshold. In addition, only 1 out of 3 factors measured by the DRS15-R demonstrated adequate internal reliability. All factors measured by the SDS, SLCS-R, and the CFQ demonstrated adequate internal reliability. Table 6.2 displays Cronbach's Alpha reliability coefficients for all measures.

Test-retest reliability

Friedman's test statistics were generally non-significant indicating that there no significant differences between HCQ-R items across sitting one and sitting two. Despite all intraclass correlations being significant (r range = 0.28 to 0.66), they were below both the conservative (> 0.80: Anastasi & Urbina, 1997; Kline, 1993) and liberal threshold (> 0.70: Vincent, 1999; see Table 6.3).

Convergent and predictive validity

Table 6.4 displays the correlations between the overlapping factors of the HCQ-R (sitting one) and the DRS15-R and SLSC-R. Despite the correlations being significant and positive, the correlations between the overlapping subscales of the HCQ-R, DRS15-R and SLSC-R were generally weak to moderate. Only the correlation between the overlapping subscale of HCQ-R confidence and SLSC-R self-liking was strong ($r > 0.70$; Fallowfield *et al.*, 2005). Findings revealed minimal association between HCQ-R factors and CFQ factors (see Table 6.4), thus providing no evidence to support the HCQ-R's predictive validity.

Table 6.1. Descriptive statistics of the HCQ-R by item across sitting one and sitting two

			Sitting one				Sitting two			
Factors		Item	Mean	*S.D.*	S ratio	K ratio	Mean	*S.D.*	S ratio	K ratio
Challenge	HCQ-R 4	I believe that change enables me to grow as a person	4.20	0.60	-2.80	4.85	4.19	0.62	-1.66	1.65
	HCQ-R 6	Changes in my daily routine encourage me to learn	3.62	0.80	-0.67	-0.77	3.57	0.80	-1.23	-0.64
	HCQ-R 11	I see challenges in my life as opportunities for me to develop as a person	4.23	0.71	-2.94	1.14	4.09	0.64	-1.31	0.87
Commitment	HCQ-R 2R	I find it difficult to stay committed to whatever I am doing	3.71	0.91	-2.84	0.17	3.70	0.84	-1.41	-0.70
	HCQ-R 8R	I lack commitment	4.14	0.79	-2.61	-0.20	4.06	0.90	-2.98	-0.42
	HCQ-R 10R	I find it difficult to stay committed to achieving my goals	3.81	0.97	-4.52	2.08	3.70	0.94	-2.64	-0.22
Control	HCQ-R 1	I can influence the path that my life takes	4.41	0.57	-2.65	2.99	4.47	0.56	-3.17	3.55
	HCQ-R 7R	My actions have little influence on my life	4.08	0.99	-6.19	4.84	4.26	1.01	-7.46	6.46
	HCQ-R 13	Events in my life are shaped by my own actions	4.21	0.61	-0.57	-0.94	4.30	0.67	-5.18	9.34
Confidence	HCQ-R 3R	I doubt whether I am a worthwhile person	3.93	0.90	-3.26	0.96	3.87	0.93	-2.11	-0.37
	HCQ-R 5	I feel that I am a person of worth	3.92	0.81	-3.20	2.32	3.86	0.80	-3.84	2.98
	HCQ-R 9	I feel that I am a truly worthwhile person	3.89	0.84	-3.02	1.64	3.81	0.92	-2.69	0.23
	HCQ-R 12R	I lack self-belief	3.71	1.02	-2.32	-1.12	3.64	1.14	-2.75	-0.79
	HCQ-R 14R	I take a negative attitude toward myself	3.71	1.05	-2.83	-0.25	3.79	1.04	-3.48	0.39

Note: S ratio = Skewness ratio; K ratio = Kurtosis ratio

Table 6.2. Descriptive statistics and Cronbach's Alpha reliability coefficients of the Revised Hardiness Confidence Questionnaire, Social Desirability Scale, Revised Dispositional Resilience Scale, Revised Self-liking/Self-competence Scale, and Coping Function Questionnaire by factor.

Questionnaires and subscales	Min	Max	Mean	*S.D.*	α
Revised Hardiness Confidence Questionnaire (sitting one)					
Challenge	3.00	5.00	4.02	0.48	0.40
Commitment	1.67	5.00	3.89	0.74	0.77
Control	2.33	5.00	4.23	0.56	0.60
Confidence	1.80	5.00	3.83	0.72	0.84
Revised Hardiness Confidence Questionnaire (sitting two)					
Challenge	2.67	5.00	4.00	0.53	0.67
Commitment	1.67	5.00	3.82	0.77	0.82
Control	1.33	5.00	4.34	0.57	0.59
Confidence	1.00	5.00	3.79	0.79	0.87
Social Desirability Scale	0	12.00	5.97	2.89	0.77
Revised Dispositional Resilience Scale					
Challenge	1.00	15.00	9.22	2.52	0.81
Commitment	0	15.00	10.33	2.14	0.65
Control	2.00	15.00	12.80	2.18	0.58
Revised Self-liking/Self-competence Scale					
Self-liking	12.00	40.00	29.22	6.06	0.90
Self-competence	10.00	40.00	26.80	4.75	0.80
Coping Function Questionnaire					
Problem-focused coping	1.00	5.00	3.13	0.99	0.84
Emotional-focused coping	1.29	4.86	3.26	0.83	0.82
Avoidance-focused coping	1.00	4.80	1.88	1.08	0.93

Table 6.3. Intraclass correlations of the Revised Hardiness Confidence Questionnaire across sitting one and sitting two.

Factors	Item	Intraclass correlation	Friedman's Chi-square
Challenge	HCQ-R 4	0.44**	0.02*
	HCQ-R 6	0.38**	0.32
	HCQ-R 11	0.28**	2.97
Commitment	HCQ-R 2R	0.52**	0.01*
	HCQ-R 8R	0.48**	0.87
	HCQ-R 10R	0.44**	1.16
Control	HCQ-R 1	0.50**	1.13
	HCQ-R 7R	0.32**	2.35
	HCQ-R 13	0.29**	1.37
Confidence	HCQ-R 3R	0.65**	0.62
	HCQ-R 5	0.66**	0.82
	HCQ-R 9	0.57**	0.97
	HCQ-R 12R	0.64**	0.57
	HCQ-R 14R	0.63**	0.80

Note: **Significance is at $p < 0.01$, * significance is at $p < 0.05$.

Table 6.4. Correlations between the Revised Hardiness Confidence Questionnaire and the Social Desirability Scale, Revised Dispositional Resilience Scale, Revised Self-liking/Self-competence Scale, and the Coping Function Questionnaire.

	Revised Hardiness Confidence Questionnaire (sitting one)			
	Challenge	Commitment	Control	Confidence
Descriptive statistics				
Social Desirability Scale	0.31**	0.21*	-0.05	0.01
Convergent validity				
Revised Dispositional Resilience Scale				
Challenge	0.22**			
Commitment		0.46**		
Control			0.46**	
Revised Self-liking/Self-competence Scale				
Self-liking				0.83**
Self-competence				0.46**
Predictive validity				
Coping Function Questionnaire				
Problem-focused coping	0.12	0.06	0.12	0.09
Emotional-focused coping	0.07	0.14	0.15	0.01
Avoidance-focused coping	-0.05	0.09	-0.06	-0.00

6.2.3 Discussion

Test-retest reliability

Findings revealed little support for the hypothesis that the HCQ-R would demonstrate acceptable levels of stability. Although Friedman's test statistics were generally non-significant indicating that there were no significant differences between HCQ-R items across sitting 1 and sitting 2, all item intraclass correlation coefficients failed to reach the conservative (r > 0.80: Anastasi & Urbina, 1997; Kline, 1993) and liberal criteria (r > 0.70: Vincent, 1999; see Table 6.3). Given that the HCQ-R purports to measure four personality traits, the preliminary test-retest coefficients obtained in Study 4 appears to question the stability of the HCQ-R. The test-retest coefficients of the HCQ-R are therefore inconsistent with those obtained in established personality inventories. For instance, MacCrae *et al.* (2011) found adequate test-retest intraclass correlations (r range = 0.70 to 0.91) for the varying facet scales of NEO personality inventories.

One possible explanation for the evidenced inadequate test-retest coefficients may be attributed to the contention that some variation in item response is inevitable given the self-report nature of psychometric instruments. Research by Nevill *et al.* (2001) stated that participants might be genuinely unclear about how they feel when responding to items. When examining the test-retest reliability of the Social Physique Anxiety Scale, Nevill and colleagues suggested that participants responding to the item "I am comfortable with the appearance of my physique/figure" might report that this is "slightly characteristic of me = 2" or "moderately characteristic of me = 3" and would have reported 2.5 if such an option was available. Nevill *et al.* posited that participants might give the item a score of 2 one day and on another day give it a score of 3, with both scores accurately representing the closest response to how an individual felt on that day. Despite both responses containing measurement error, both responses are valid assessments of the target construct (Nevill *et al.*). Subsequently, one may tentatively argue that participants' clarity of their feelings when completing the HCQ-R could have contributed to the inadequate test-retest reliability coefficients obtained in Study 4. Given the importance placed on examining test-retest reliability in the early stages of scale development particularly when assessing personality traits (e.g., McCrae *et al.*, 2011; Nevill *et al*; Watson, 2004), one may question the ability of the HCQ-R to yield stable scores over time. Further research is therefore warranted to examine the stability of the HCQ-R.

Convergent validity

The findings appear to provide little support for the hypothesis that the HCQ-R would exhibit acceptable convergent validity in that significant yet weak to moderate correlations were generally obtained between the respective overlapping subscales of the HCQ-R and the DRS15-R and the SLCS-R (see Table 6.4). Consequently, using the recommendations outlined by

MacKenzie *et al.* (2011) and Marsh (2007; $r > 0.70$), the findings of Study 4 provide little support for the convergent validity of the HCQ-R (see Table 6.4). However, inspection of the correlation values used to support the convergent validity of instruments in the MT literature and beyond have been relatively less stringent than the parameters suggested by MacKenzie *et al.* and Marsh. For instance, when developing their Cricket Mental Toughness Inventory, Gucciardi and Gordon (2009) correlated the CMTI with dispositional flow, burnout, hardiness, and dispositional resilience. Correlations between overlapping subscales of the CMTI (desire to achieve, attentional control) and hardiness (commitment, control) were positive and significant yet only weak ($r = 0.23$, 0.28, respectively) according to the guidelines of Fallowfield *et al.* (2005). In addition, Gucciardi *et al.* (2009) found similar findings when assessing the convergent validity of the Australian football Mental Toughness Inventory in that correlations between overlapping subscales of the AfMTI (thrive through challenge, desire success) and hardiness (challenge, commitment) were positive and significant yet could be considered weak ($r = 0.29$, 0.23, respectively).

Furthermore, researchers developing measures beyond MT have also used correlations which are below MacKenzie *et al.'s* (2011) and Marsh's (2007) recommendations to demonstrate convergent validity. For example, Zourbanos *et al.* (2009) correlated their Automatic Self-talk Questionnaire for Sports (ASTQS) with anxiety, affect, and performance strategy usage. Correlations of overlapping subscales revealed a positive and significant ($p < 0.05$) yet weak relationships between the ASTQS 'psych up' factor and the 'vigour' facet of affect ($r = 0.26$), and ASTQS 'irrelevant thoughts' with the 'concentration disruption' facet of anxiety ($r = 0.23$). Given that the negative ASTQS subscales had positive and significant correlations with negatively valenced and negative correlations with positively valenced subscales from other instruments, Zourbanos *et al.* argued that the pattern of correlations provides adequate convergent validity for the ASTQS.

The use of correlation patterns (i.e., significance, hypothesised direction) has been used by many researchers (e.g., Freeman *et al.*, 2011; Lonsdale *et al.*, 2008; Williams & Cummings, 2011) in the scale development literature to demonstrate convergent validity. Although the convergent validities in Study 4 do not meet MacKenzie *et al.'s* (2011) and Marsh's (2007) criteria, the weak to moderate, significant, and positive correlations between the overlapping subscales of the HCQ-R and the DRS15-R and SLSC-R are comparable with the data used in recent research to demonstrate convergent validity. Furthermore, Kline (2005) argues that correlations ranging from 0.30 to 0.50 are indicative of nomological validity. Consequently, when using the patterns of correlations and more liberal thresholds as a guide, the findings of Study 4 provide preliminary support for the convergent validity of the HCQ-R.

One possible explanation as to why the correlations between the overlapping subscales of the HCQ-R, DRS15-R and SLSC-R did not conform to the conservative criteria could be due to the inherent variability in responses when assessing matching traits by two different instruments (Marsh *et al.,* 2010). Marsh and colleagues argue that the extent to which matching traits actually match is likely to vary when utilising different instruments which have been independently developed by different researchers. Inspection of the response formats used by the HCQ-R, DRS15-R, and SLSC-R may provide an explanation for the convergent validities obtained in Study 4. Although the response format of the HCQ-R and SLSC-R to measure confidence were identical (i.e., anchored by (1) strongly disagree to (5) strongly agree), the response format of the DRS15-R is slightly different (i.e., anchored by (0) not at all true to (3) completely true) to the HCQ-R. Consequently, one may tentatively suggest that the response format of the respective questionnaires used to assess convergent validity may have contributed towards the weak convergent validities obtained for the challenge, commitment, and control factors.

Predictive validity

The findings provided very limited support for the hypothesis that HCQ-R scores will be positively and significantly correlated to problem- and emotion-focused coping strategies, and significantly and negatively correlated to the use of avoidance-focused strategies. Correlations revealed very little association between HCQ-R scores and CFQ scores (see Table 6.4). Using the guidelines proposed by Fallowfield *et al.* (2005), the correlations obtained indicate no relationship between HCQ-R scores and CFQ scores. Consequently, no regression analyses were conducted. The predictive validity findings of Investigation 1 of Study 4 are therefore inconsistent with the theoretical predictions of hardiness (Kobasa, 1979; Maddi & Kobasa, 1984) whereby high levels of challenge, commitment, and control are thought to be associated with problem- and emotion-focused coping and negatively associated with avoidance-focused coping. Moreover, the findings are in contrast to previous research which has generally supported the contention that challenge, commitment, and control are positively related with problem- and emotion-focused coping and negatively related to avoidance-focused coping (e.g., Florian *et al.,* 1995; Klag & Bradley, 2004; Maddi, 1999; Maddi & Hightower, 1999; Wadey *et al.*, 2012; Williams *et al.*, 1992).

One possible reason to explain the inconsistent predictive validity findings of Investigation 1 could be due to the methodological limitations associated with using questionnaires to measure previous coping behaviours. Research (e.g., Ptacek, Smith, Espe, & Rafferty, 1994; Smith, Leffingwell, & Ptacek, 1999; Stone *et al.,* 1998; Todd, Tennen, Carney, Armeli, & Affleck, 2004) examining the temporal accuracy of questionnaires has generally supported the contention that with an increasing passage of time people provide less accurate coping accounts. For example, Stone *et al.* (1998) compared momentary assessment reports taken at multiple time points across 48 hours and retrospective reports of coping taken directly after the final momentary assessment and found that participants retrospectively over reported behavioural strategies and under reported cognitive strategies. In addition, 30% of participants failed to retrospectively report items that they had reported on the momentary assessments and 30% of participants retrospectively reported items that were not reported on the momentary assessments despite the same questionnaires being used. These findings would suggest that the validity of the coping accounts yielded in Study 4 are questionable given that the CFQ required participants to retrospectively recall coping accounts of the most stressful sporting event within the previous 12 months. Due to the inherent limitations of biographical memory, participants completing the CFQ may have forgotten, under reported, or over reported when retrospectively recalling the coping strategies used to deal with their self-reported stressor (Folkman & Moskowitz, 2004). This could have contributed towards the unexpected predictive validity findings in Investigation 1. Future research using coping as a variable to assess an instrument's predictive validity should therefore assess coping strategies as close to the stressful event as possible in an effort to overcome the discussed limitations of retrospective recall.

Another possible reason to explain the lack of predictive validity of the HCQ-R could be attributable to its psychometric properties. Despite the HCQ-R being developed in accordance with recommendations outlined in the scale develop literature (e.g., MacKenzie *et al.*, 2011; Marsh, 2002), support for its factorial validity was only obtained when using more liberal model fit thresholds. Furthermore, the internal reliability estimates obtained for the HCQ-R challenge and control subscales in Investigation 1 suggest that there may be problems with the extent to which items reflecting the same construct yield similar results (Nunnally & Bernstein, 1994). However, caution should be taken when interpreting the internal consistency findings of the HCQ-R given the multidimensionality of its underpinning model (Schmitt, 1996), the relatively small number of items making up the instrument and the relatively small sample size used in Investigation 1 (Gignac, 2009). Given the lack of predictive validity of the HCQ-R obtained in Investigation 1, research is required to overcome the methodological limitations inherent in retrospectively recalling coping behaviours and examine the extent to which the instrument could predict real-time pre-competition cognitions.

6.3 Investigation 2

In an effort to further examine the HCQ-R's predictive validity in sport, Investigation 2 presents an examination of the relationships between HCQ-R scores and athletes' challenge and threat appraisals and perceptions of pre-competitive state anxiety prior to sport competition. A number of hypothesises were formulated. It was hypothesised that HCQ-R scores would be negatively and significantly correlated to perceived levels of pre-competitive anxiety intensity and perceived frequency of anxiety experienced; positively and significantly correlated with more facilitative perceptions of pre-competitive anxiety; positively and significantly correlated with perceived intensity, positive interpretations, and frequency of self-confidence experienced; positively and significantly correlated with challenge appraisals; and negatively and significantly correlated with threat appraisals. It was also hypothesised that HCQ-R scores would significantly predict the use of challenge and threat appraisals and pre-competitive state anxiety. Specifically, it was hypothesised that HCQ-R confidence would predict cognitive and somatic anxiety intensity and frequency, and self-confidence intensity, interpretation, and frequency. For cognitive and somatic anxiety interpretations, it was hypothesised that HCQ-R control would make unique contributions to the regression models. It was also hypothesised that HCQ-R challenge would predict challenge and threat appraisals.

6.3.1 Method

Participants

As stated in Investigation 1, the minimum sample required to conduct multiple regression analyses on the HCQ-R is 82 participants (Green, 1991). One hundred and forty one competitive student athletes from a University in Southern England were recruited to participate in this study. Thirty-seven athletes were substitutes for their respective competitions and were subsequently removed from the data analysis. This resulted in a sample of 104 competitive student athletes. Participants consisted of 62 males and 42 females (*M* age = 20.36 years, *S.D.* = 1.69) with a mean of 10.11 years (*S.D.* = 3.64) competitive playing experience in their primary sport. The majority of participants (88.5%) were White British, with the remaining participants (*n* = 12) consisting of a variety of ethnic backgrounds including White other (*n* = 3), Mixed other (*n* = 2), White Irish (*n* = 1), Mixed White and Black Caribbean (*n* = 1), Mixed White and Black African (*n* = 1), Asian Bangladeshi (*n* = 1), Black or Black British African (*n* = 1), Chinese (*n* = 1), and any other ethnic background (*n* = 1). Participants competed in a variety of sports including Football (*n* = 37), Netball (*n* = 23), Rugby (*n* = 10), Basketball (*n* = 10), Badminton (*n* = 5), Hockey (*n* = 7), and Volleyball (*n* = 7). The highest level of primary sport participation ranged from Intervarsity (*n* = 16) through Club (*n* = 35), County (*n* = 28), Regional (*n* = 13), National (*n* = 11), and International (*n* =1) level. All participants were currently representing their University sports teams in the British University and Colleges league.

Measures

Revised Hardiness Confidence Questionnaire

See Investigation 1 of Study 4 for details of the HCQ-R.

Challenge and threat construal measure

An adapted 6-item version of McGregor and Elliot's (2002) challenge and threat construal measure was used to assess participants' pre-competition appraisals (see Appendix 4.11). Definitions of appraisals were provided to clarify how individuals might respond to competition. Item wording was adapted to tap into pre-competition appraisals. The three items reflecting a challenge appraisal included "I view today's match as a challenge", "I am looking forward to being positively challenged in today's match", and "I think today's match represents a positive challenge to me". The three items representative of a threat appraisal included "I view today's match as a threat", "I think today's match could be threatening to me", and "I think today's match represents a threat to me". Participants were required to respond using a 7-point Likert scale ranging from "not at all true of me" (1) to "very true of me" (7). Two subscale scores for challenge and threat were obtained by calculating the mean score (see Appendix 4.12 for the challenge and threat construal measure scoring key). The challenge and threat construal measure has yielded adequate internal consistency (challenge: $\alpha = 0.78$; threat: $\alpha = 0.73$) and factorial validity (challenge: RCFI = 0.99, SRMR = 0.03, RMSEA 0.04; threat: RCFI = 0.98, SRMR = 0.03, RMSEA 0.09) in research involving athletes similar in age to the athletes used in this study (Aide, Duda, & Ntoumanis, 2008).

Immediate Anxiety Measurement Scale

The Immediate Anxiety Measurement Scale (IAMS: Thomas, Hanton, & Jones, 2002) was used to assess the intensity, directional, and frequency perceptions of pre-competitive anxiety and self-confidence experienced by participants (see Appendix 4.13). The IAMS measures cognitive anxiety intensity (CAI), cognitive anxiety direction (CAD), cognitive anxiety frequency (CAF), somatic anxiety intensity (SAI), somatic anxiety direction (SAD), somatic anxiety frequency (SAF), self-confidence intensity (SCI), self-confidence direction (SCD), and self-confidence frequency (SCF). Participants rated their intensity perceptions on a scale ranging from "not at all" (1) to "extremely" (7), direction on a scale ranging from "very debilitative/negative" (-3) to "very facilitative/positive" (+3) with 0 indicating "unimportant", and frequency of intrusions on a scale ranging from "not at all" (1) to "all the time" (7). Prior to responding, participants were required to read definitions of each construct to enable individuals to fully understand their respective meaning (see Appendix 4.13 for a full list of definitions). Research by Thomas *et al.* has provided support for the convergent validity of the IAMS when correlated with the Competitive State Anxiety Inventory-2 (Martens, Burton, Vealey, Bump, & Smith, 1990).

Social Desirability Scale

See Study 1 for details of the SDS.

Match Importance Scale

The match importance scale was developed to assess the importance placed on the upcoming competition by participants. Participants rated the importance of the competition on an 11-point Likert scale ranging from “not important” (1) to “important” (11).

Procedure

Participants were provided with a cover story and informed that the study was investigating the psychological characteristics and thought processes of competitive athletes which required them to complete a series of questionnaires prior to training and competition. Prior to a standard training session, participants were provided with an information sheet, consent form, demographic questionnaire, the HCQ-R, and the SDS (see Appendix 4.14 for the information sheet, consent form, and demographic questionnaire). In order to overcome the inadequacies of using cross-sectional designs when assessing an instrument’s predictive validity (Gignac, 2009), participants were required to complete the challenge and threat construal measure, the IAMS, and the match importance scale prior to competition. All questionnaires were completed in a seated environment (e.g., classroom). The duration between the first data collection (pre-training) and the second data collection (pre-competition) ranged from 5 to 9 days. Pre-competition questionnaires were administered between 30 minutes and two hours before competition and were completed in the presence of the author so that any questions could be answered. Participants were recruited over a period of six weeks. The pre-training questionnaires took approximately 5-8 minutes to complete and the pre-competition questionnaires took approximately 2-5 minutes to complete. Upon completing the pre-competition questionnaires, participants were thanked for their participation and received a verbal and written debrief explaining the nature of the study (see Appendix 4.15 for written debrief).

Data analyses

Data screening

Analysis of skewness and kurtosis ratios for HCQ-R scores showed that the data were marginally nonnormally distributed (see Table 6.5), with negative skewness showing that the data are distributed towards the upper end of the HCQ-R scale. Although inspection of HCQ-R box-plots identified 50 outliers relating to 34 participants (see Appendix 4.16), all participants were taken forward for subsequent regression analysis given the absence of any theoretical explanation to underpin their removal (Tabachnick & Fidell, 2001). Analysis of skewness and kurtosis ratios for the challenge and threat construal measure and the IAMS generally showed that the data were normally distributed except for the challenge subscale (see Appendix 4.17). Box-plots for the challenge subscale of the challenge and threat construal measure revealed 14 outliers relating to 10 participants (see Appendix 4.18). Given the lack of theoretical explanation to underpin their removal, all participants were taken forward for regression analyses.

Scatterplots between the respective subscales of the HCQ-R and the challenge and threat construal measure and the IAMS generally provided support for the assumption of linearity (see Appendix 4.19 for scatter diagrams of significant relationships between independent and dependent variables).

Regression analyses

Correlations were conducted to assess the relationships between the HCQ-R and the respective pre-competition measures. Hierarchical regression analyses were used to examine the main effects of MT upon challenge/threat appraisals and anxiety and self-confidence perceptions experienced prior to competition. In the hierarchical regression analyses, the dependent variables were the respective subscales of the challenge and threat construal measure and the IAMS. Given that some of the dependent variables showed little association to the respective subcomponents of MT, regression analyses were only conducted on those dependent variables which showed a significant relationship (see Table 6.7). This resulted in regression analyses being conducted between the subcomponents of the HCQ-R and CAD, SAI, SAD, SAF, SCI, SCD, SCF, challenge, and threat.

The challenge, commitment, control, and confidence factors were entered in blocks depending on their theoretical association with the dependent variable. For CAD and SAD, control was entered in block one based on its theoretical importance to competitive state anxiety interpretations. According to Jones' (1995) Control Model of Debilitative and Facilitative Competitive State Anxiety, control is a key determinant of anxiety interpretations whereby high levels of control result in positive anxiety interpretations and low levels of control result in negative interpretations. Given the theoretical associations between challenge appraisals (e.g., Jones *et al.*, 2009) and confidence (e.g., Martens *et al.*, 1990) with competitive state anxiety perceptions, challenge and confidence were entered in block two, with commitment being entered in block three. For SAI and SAF, confidence was entered in block one on the grounds that it is theorised to be a key determinant of competitive state anxiety intensity and frequency perceptions (e.g., Martens *et al.*), with challenge, commitment, and control being entered in block two. For SCI, SCD, and SCF, confidence was entered in block one based on its direct theoretical association with competitive state self-confidence perceptions, with challenge, commitment, and control being entered in block two. For pre-competitive challenge and threat states, HCQ-R challenge was entered in block one based on its direct theoretical association with challenge and threat appraisals. Given the importance placed on control and confidence in Jones *et al.'s* Theory of Challenge and Threat States in Athletes, control and confidence were entered in block two, with commitment being entered in block three. The explained variance (R^2) in the dependent variable and the sign of the regression coefficients (b) were assessed. A significance level of $p < 0.05$ was used for all analyses.

6.3.2 Results

Descriptive statistics

Table 6.5 displays the descriptive statistics for the HCQ-R item scores. Descriptive statistics of the HCQ-R, SDS, the challenge and threat construal measure, and the IAMS by factor are displayed in Table 6.6. Findings suggest that socially desirable responding had little impact on the scores obtained for the HCQ-R in that correlations with the SDS were weak (see Table 6.7). Match importance ratings were generally very high (*M* rating = 9.73, *S.D.* = 1.57) suggesting that the matches used in Investigation 2 were suitable to elicit pre-competitive challenge and threat appraisals and pre-competitive anxiety and self-confidence symptoms. Inspection of Mahalanobis distance statistics for HCQ-R scores showed that the maximum value (13.96) did not exceed Tabachnick and Fidell's (2001) critical chi square ratio (18.47) for models with four independent variables, thus revealing no multivariate outliers (see Appendix 4.20). Scatterplots between the regression standardised residual and the regression standardised predicted value of each dependent variable generally showed support for the assumption of homeoscedasticity (see Appendix 4.21). Inspection of the Variance Inflation Factor (VIF) and its related tolerance value showed little evidence of multicollinearity as the largest VIF values did not exceed 10 (Myers (1990) and the tolerance values generally exceeded 0.20 (Menard, 1995; see Appendix 4.22).

Regression analyses

Correlation analyses showed a number of dependent variables to be significantly related to the constructs of the HCQ-R (see Table 6.7). The results of the respective hierarchical regression analyses are shown in Table 6.8. All four HCQ-R constructs significantly predicted SAD (R^2 = 0.11, $p < 0.05$), SAF (R^2 = 0.17, $p < 0.01$), SCI (R^2 = 0.15, $p < 0.01$), SCD (R^2 = 0.14, $p < 0.01$), and challenge (R^2 = 0.24, $p < 0.01$). The effects on SAD and challenge were primarily attributable to HCQ-R challenge (b = 1.00, β = 0.37, $p < 0.05$; b = 0.61, β = 0.34, $p < 0.01$, respectively), whereas the effects on SCI and SCD were primarily attributable to HCQ-R confidence (b = 0.72, β = 0.30, $p < 0.01$; b = 0.80, β = 0.27, $p < 0.05$, respectively). For SAF, HCQ-R confidence (b = -1.02, β = -0.33, $p < 0.01$) and commitment (b = -0.58, β = -0.20, $p < 0.05$) were found to make a unique contribution to the model. Although HCQ-R challenge, control, and confidence provided the only statistically significant predictor of threat (R^2 = 0.08, $p < 0.05$), all coefficients failed to reach significance. Only HCQ-R confidence was found to be a significant predictor of SAI (R^2 = 0.06, $p < 0.05$). The regression models for CAD and SCF were not found to be statistically significant at any step examined (see Table 6.8).

Table 6.5. Descriptive statistics of the HCQ-R by item.

Factors		Item	Mean	*S.D.*	S ratio	K ratio
Challenge	HCQ-R 4	I believe that change enables me to grow as a person	3.88	0.73	-3.13	3.89
	HCQ-R 6	Changes in my daily routine encourage me to learn	3.67	0.73	-2.64	-0.64
	HCQ-R 11	I see challenges in my life as opportunities for me to develop as a person	4.05	0.73	-3.55	2.77
Commitment	HCQ-R 2R	I find it difficult to stay committed to whatever I am doing	4.15	0.55	0.26	0.15
	HCQ-R 8R	I lack commitment	4.24	0.84	-7.08	9.42
	HCQ-R 10R	I find it difficult to stay committed to achieving my goals	4.12	0.69	-3.74	6.35
Control	HCQ-R 1	I can influence the path that my life takes	4.30	0.59	-1.99	2.28
	HCQ-R 7R	My actions have little influence on my life	4.32	0.69	-5.24	8.43
	HCQ-R 13	Events in my life are shaped by my own actions	3.96	0.64	-1.80	1.91
Confidence	HCQ-R 3R	I doubt whether I am a worthwhile person	4.35	0.69	-2.48	1.63
	HCQ-R 5	I feel that I am a person of worth	4.10	0.69	-1.30	-0.47
	HCQ-R 9	I feel that I am a truly worthwhile person	4.07	0.77	-2.16	-0.13
	HCQ-R 12R	I lack self-belief	4.18	0.76	-3.64	3.32
	HCQ-R 14R	I take a negative attitude toward myself	4.13	0.81	-4.78	5.56

Note: S ratio = Skewness ratio; K ratio = Kurtosis ratio.

Table 6.6. Descriptive statistics of the Revised Hardiness Confidence Questionnaire, Social Desirability Scale, Immediate Anxiety Measurement Scale, and challenge and threat construal measure.

Questionnaires and subscales	Min	Max	Mean	*S.D.*
Revised Hardiness Confidence Questionnaire				
Challenge	2.67	5.00	3.87	0.52
Commitment	3.00	5.00	4.17	0.52
Control	3.33	5.00	4.19	0.42
Confidence	2.60	5.00	4.16	0.48
Social Desirability Scale	1.00	11.00	5.79	2.54
Immediate Anxiety Measurement Scale				
CAI	1.00	6.00	3.98	1.25
CAD	-2.00	3.00	1.00	1.25
CAF	1.00	7.00	3.89	1.34
SAI	1.00	7.00	3.35	1.47
SAD	-2.00	3.00	0.54	1.39
SAF	1.00	7.00	3.72	1.50
SCI	2.00	7.00	4.99	1.15
SCD	-2.00	3.00	1.60	1.41
SCF	1.00	7.00	4.71	1.23
Challenge and threat construal measure				
Challenge	2.00	7.00	5.72	0.94
Threat	1.00	6.33	2.97	1.20

Note: CAI = cognitive anxiety intensity; CAD = cognitive anxiety direction; CAF = cognitive anxiety frequency; SAI = somatic anxiety intensity; SAD = somatic anxiety direction; SAF = somatic anxiety frequency; SCI = self-confidence intensity; self-confidence direction = SCD; SCF = self-confidence frequency.

Table 6.7. Correlations between the Revised Hardiness Confidence Questionnaire and the Social Desirability Scale, Immediate Anxiety Measurement Scale, and challenge and threat construal measure.

	Revised Hardiness Confidence Questionnaire			
	Challenge	Commitment	Control	Confidence
Descriptive statistics				
Social Desirability Scale	0.14	0.08	0.01	- 0.07
Predictive validity				
Immediate Anxiety Measurement Scale				
CAI	0.02	0.02	- 0.18	- 0.04
CAD	0.18	0	0.17	0.23*
CAF	0.09	- 0.15	- 0.02	- 0.08
SAI	- 0.14	- 0.18	- 0.09	- 0.28**
SAD	0.26**	- 0.03	- 0.04	0.03
SAF	- 0.15	- 0.31**	- 0.02	- 0.34**
SCI	0.22*	0.13	0.23*	0.38**
SCD	0.28**	0.05	0.18	0.34**
SCF	0	- 0.02	0.20*	0.19
Challenge and threat construal measure				
Challenge	0.45**	0.25*	0.30**	0.45**
Threat	- 0.15	- 0.12	- 0.24*	- 0.24*

Note: **Significance is at $p < 0.01$, *significance is at $p < 0.05$.

Table 6.8. Hierarchical regression analyses: Effects of HCQ-R dimensions upon pre-competitive perceptions of anxiety, self-confidence, and challenge and threat appraisals.

Dependent variable	Step	R^2	ΔR^2	*b* (*S.E.*)	β
CAD	1: Control	0.04		0.56 (0.29)	0.19
	2: Control	0.07	0.03	0.33 (0.32)	0.11
	Challenge			0.11 (0.26)	0.05
	Confidence			0.44 (0.29)	0.17
	3: Control	0.09	0.02	0.31 (0.31)	0.10
	Challenge			0.18 (0.26)	0.07
	Confidence			0.55 (0.30)	0.21
	Commitment			-0.38 (0.25)	-0.16
SAI	1: Confidence	0.06*		-0.77 (0.29)*	-0.26
	2: Confidence	0.07	0.01	-0.66 (0.36)	-0.22
	Challenge			0.01 (0.31)	0
	Commitment			-0.23 (0.30)	-0.08
	Control			-0.08 (0.37)	-0.02
SAD	1: Control	0		-0.17 (0.33)	-0.05
	2: Control	0.10*	0.10	-0.39 (0.35)	-0.12
	Challenge			0.93 (0.29)*	0.35
	Confidence			-0.29 (0.32)	-0.10
	3: Control	0.11*	0.01	-0.40 (0.34)	-0.12
	Challenge			1.00 (0.29)*	0.37
	Confidence			-0.19 (0.33)	-0.06
	Commitment			-0.36 (0.27)	-0.13
SAF	1: Confidence	0.13**		-1.14 (0.29)**	-0.37
	2: Confidence	0.17**	0.04	-1.02 (0.34)**	-0.33
	Challenge			0.01 (0.30)	0
	Commitment			-0.58 (0.29)*	-0.20
	Control			0.27 (0.36)	0.07
SCI	1: Confidence	0.13**		0.88 (0.22)**	0.37
	2: Confidence	0.15**	0.02	0.72 (0.27)**	0.30
	Challenge			0.21 (0.24)	1.00
	Commitment			-0.08 (0.22)	-0.04
	Control			0.24 (0.28)	0.09

Table 6.8. *Cont.*

Dependent variable	Step	R^2	ΔR^2	*b* (*S.E.*)	β
SCD	1: Confidence	0.09**		0.88 (0.29)**	0.30
	2: Confidence	0.14**	0.05	0.80 (0.33)*	0.27
	Challenge			0.51 (0.29)	0.19
	Commitment			-0.50 (0.28)	-0.18
	Control			0.07 (0.35)	0.02
SCF	1: Confidence	0.04		0.49 (0.25)	0.19
	2: Confidence	0.06	0.02	0.51 (0.30)	0.20
	Challenge			-0.07 (0.27)	-0.03
	Commitment			-0.27 (0.25)	-0.11
	Control			0.34 (0.31)	0.11
Challenge	1: Challenge	0.19**		0.80 (0.16)**	0.44
	2: Challenge	0.24**	0.05	0.61 (0.18)**	0.34
	Confidence			0.28 (0.20)	0.14
	Control			0.29 (0.22)	0.13
	3: Challenge	0.24**	0	0.61 (0.18)**	0.34
	Confidence			0.28 (0.21)	0.14
	Control			0.29 (0.22)	0.13
	Commitment			-0.01 (0.17)	0
Threat	1: Challenge	0.04		-0.44 (0.23)	-0.19
	2: Challenge	0.08*	0.04	-0.23 (0.25)	-0.10
	Confidence			-0.53 (0.30)	-0.18
	Control			-0.20 (0.28)	-0.08
	3: Challenge	0.08	0	-0.24 (0.26)	-0.10
	Confidence			-0.53 (0.30)	-0.18
	Control			-0.20 (0.29)	-0.08
	Commitment			0 (0.24)	0

6.3.3 Discussion

The findings provide support for the hypothesis that HCQ-R scores would predict pre-competitive challenge appraisals (see Table 6.7 & Table 6.8). The regression analyses appear to provide a clear hierarchy in regards to the relative contributions of the traits which predict challenge appraisals in athletes. HCQ-R challenge scores accounted for 19% of variance in pre-competitive challenge appraisals which represented the largest amount of explained variance by a single HCQ-R construct. This finding demonstrates the relative importance of dispositional challenge when predicting challenge states in athletes and suggests that athletes who embrace change in their lives and view changes as interesting incentives to growth (Kobasa *et al.*, 1982) are more likely to appraise competition with a challenge state. This finding is not surprising given the theoretical association between dispositional challenge and challenge states. The findings also highlight the relative contribution of HCQ-R confidence and control in predicting challenge states in athletes in that a significant change in combined variance (5%) was explained above and beyond the variance explained by HCQ-R challenge (see Table 6.8). This finding demonstrates that although dispositional challenge provided the most salient predictor of challenge appraisals, HCQ-R confidence and control also provided meaningful information to explain its variance. These findings are consistent with the theoretical predictions of Jones *et al.'s* (2009) Theory of Challenge and Threat States in Athletes whereby high levels of confidence and control are posited to promote challenge states in athletes. According to Jones and colleagues (2009), individuals who believe they have the necessary skills to cope with the demands of the situation are more likely to experience a pre-competitive challenge state. Furthermore, individuals who believe they have sufficient control (over the environment) to display those skills will be more likely to experience challenge states (Jones *et al.*). Notwithstanding, the effects on challenge appraisals were primarily attributable to HCQ-R challenge ($b = 0.61$, $\beta = 0.34$, $p < 0.01$), which in turn provides support for the hypothesis that HCQ-R challenge would provide the most meaningful prediction of pre-competitive challenge appraisals.

Support was found for the hypothesis that HCQ-R scores would predict threat appraisals. In contrast to pre-competitive challenge appraisals, the effects on pre-competitive threat appraisals are less clear. Although HCQ-R challenge explained 4% of variance in pre-competitive threat appraisals, this prediction was not statistically significant (see Table 6.8). This finding is therefore inconsistent with the model examining the effects on pre-competitive challenge appraisals whereby HCQ-R challenge significantly explained the majority of variance in pre-competitive challenge states. However, the addition of HCQ-R confidence and control resulted in a significant prediction of pre-competitive threat appraisals, explaining 8% of the variance. These findings therefore indicate that although the predictive capacity of dispositional challenge appears to be limited when assessed independently, the combined contributions of challenge, confidence, and control provide a meaningful prediction of threat states in athletes. Subsequently, athletes high in dispositional challenge, confidence, and control are less likely to approach competition with a threat state. The effects on pre-competitive threat appraisals further supports the theoretical predictions of Jones *et al.* (2009) whereby confidence and control are thought to be key determinants in the appraisal of stress. However, findings did not identify any HCQ-R constructs as being unique contributors to the prediction of pre-competitive threat states, which in turn provides no support for the hypothesis that effects on pre-competitive threat states would be primarily attributable to HCQ-R challenge.

Findings revealed that HCQ-R commitment did not contribute to the explained variance of pre-competitive challenge and threat appraisals (see Table 6.8). The absence of a meaningful contribution of dispositional commitment is therefore inconsistent with the hardiness literature (e.g., Kobasa, 1979, Maddi & Kobasa, 1984) whereby a combination of challenge, commitment, and control are theorised to enable an individual to cognitively restructure a stressful event so it appears less threatening. Furthermore, the findings of Investigation 2 appear to be somewhat inconsistent with the research by Weibe (1991) whereby individuals with high levels of hardiness (combined score of challenge, commitment, and control) perceived a stressor as less threatening than low hardy individuals. Inspection of the theoretical determinants of challenge and threat appraisals outlined in the sport and exercise psychology literature may provide an explanation for the apparent redundancy of dispositional commitment. Although dispositional challenge shares an intuitive association with challenge and threat states and theoretical models have outlined the relative importance of control and confidence (e.g., Jones *et al.*, 2009), the impact of commitment on stress appraisal has received little attention in the research literature. Consequently, one could argue that commitment is theoretically less important in predicting challenge and threat states in athletes, which in turn, provides an explanation for the findings of Investigation 2.

The findings therefore indicate support for HCQ-R scores to predict pre-competitive challenge and threat states in athletes, with superior support being found for the prediction of challenge states. Research examining challenge and threat states in sport has generally found that athletes who adopt a challenge state perform better than individuals who adopt a threat state (e.g., Blascovich, Seery, Mugidge, Norris, & Weisbuch, 2004; Moore, Vine, Wilson, & Freeman, 2012). Consequently, one may tentatively suggest that HCQ-R scores may not only provide some meaningful information to predict pre-competitive challenge and threat appraisals, but may have the potential to predict subsequent performance. Future research is therefore required to examine the effects of HCQ-R scores on sporting performance to further shed light on the instrument's predictive validity.

The findings provided support for the hypothesis that HCQ-R scores would predict perceived intensity, frequency, and facilitative interpretations of self-confidence. Although the respective regression models for SCF were not statistically significant, all four HCQ-R constructs significantly predicted SCI and SCD, accounting for 15% and 14% of explained variance, respectively. Given that the effects on SCI and SCD were primarily attributable to HCQ-R confidence, the findings suggest that athletes high in dispositional confidence were more likely to experience high levels and more positive interpretations of pre-competitive self-confidence. Research (e.g., Parfitt & Pates, 1999; Smith, Bellamy, Collins, & Newell, 2001; Thelwell & Maynard, 1998) investigating the impact of self-confidence on sporting performance has generally shown that individuals high in self-confidence perform better than individuals low in self-confidence. Furthermore, Kais and Raudsepp (2004) examined the relative contribution of self-confidence in the anxiety-performance relationship and found self-confidence to be superior in predicting performance when compared to perceived interpretations of cognitive and somatic anxiety. Accordingly, one may tentatively argue that the findings of Investigation 2 suggest that individuals scoring highly on HCQ-R confidence might perform better than individuals scoring low on this construct. Given that Study 4 did not measure performance, future research is required to examine the effects of HCQ-R confidence on subsequent performance outcomes.

Support was found for the hypothesis that HCQ-R scores would predict more facilitative perceptions of pre-competitive somatic anxiety. Although HCQ-R control provided no information to explain SAD ($R^2 = 0$, $p > 0.05$), the addition of HCQ-R challenge, confidence, and commitment provided a significant prediction, accounting for 11% of variance explained. Specifically, HCQ-R challenge was found to make a unique contribution in explaining the variance in SAD. These findings therefore suggest that athletes high in dispositional challenge were more likely to interpret somatic anxiety as facilitative rather than debilitative to performance. Subsequently, no support was found for the hypothesis that the effects on pre-competitive anxiety interpretations would be primarily attributable to HCQ-R control. This finding is therefore inconsistent with the theoretical predictions of Jones' (1995) Control Model of Debilitative and Facilitative Competitive State Anxiety which argues control to be a central determinant of whether anxiety is interpreted as helpful or unhelpful to performance.

The findings provided support for the hypothesis that HCQ-R scores would be negatively related to perceived levels and frequency of pre-competitive anxiety. However, given that regression analyses showed that only HCQ-R confidence significantly predicted SAI, accounting for 6% of the variance, caution is needed when interpreting these findings. Although all four HCQ-R constructs significantly predicted SAF, the effects were primarily attributable to HCQ-R confidence and commitment (see Table 6.8). The effects on SAI and SAF generally suggest that athletes high in dispositional confidence are less likely to experience high levels and frequent symptoms of somatic anxiety prior to competition. This finding is consistent with the notion that confidence may buffer against the potential adverse effects of competitive state anxiety (e.g., Hardy, 1996; Hardy, Woodman, & Carrington, 2004). In contrast, the findings of Investigation 2 provided no support for the hypothesis that HCQ-R scores would significantly predict pre-competitive cognitive anxiety intensity, interpretation, and frequency. Despite HCQ-R confidence being positively related to CAD (see Table 6.7), the regression model was not statistically significant. Given the minimal and non-significant relationships between HCQ-R scores and CAI and CAF, no regression analyses were performed.

The findings of Investigation 2 of Study 4 provide preliminary support for the HCQ-R's predictive validity. These findings therefore appear to indicate that the traits thought to underpin MT (challenge, commitment, control, confidence) provide meaningful information to predict pre-competitive cognitions in sport. Although the relationships are relatively weak (see Table 6.7), the correlation values are consistent with previous research which have examined an instrument's nomological validity both in MT (e.g., Gucciardi & Gordon, 2009: r range = -0.01 to 0.54) and beyond (e.g., Freeman *et al.*, 2011: r range = -0.37 to 0.44). Notwithstanding, the most salient reason to explain the relatively low correlations obtained in Investigation 2 could be attributed to the extent to which traits actually predict cognition and behaviour. Despite Clough *et al.* (2002) and Earle (2006) proposing that MT is underpinned by a combination of traits which predispose an individual to think and behave in a certain way, predispositions do not mean that an individual will always think and behave in the same way regardless of the situation (Aidman & Schofield, 2004). It is therefore important to highlight the role of the environment when examining the extent to which MT traits predict pre-competition cognitions. Specifically, the importance of both dispositions and the environment has led many researchers to examine cognition and behaviour from an interactional approach whereby personal traits and situational factors are thought to collectively influence how an individual thinks and behaves (Aidman & Schofield, 2004). Although match importance findings provided information to suggest that participants deemed the match to be highly important, other situational factors may have influenced the extent to which MT traits predicted pre-competition cognitions in Investigation 2. For instance, an athlete with high trait confidence will not necessarily be confident in all situations. If an athlete perceives the opponent as having superior skills to them or has a history of losing against the opponent, trait confidence may interact with these situational factors and influence the athletes perceived level of pre-competitive state confidence. This in turn, could influence the predictive validity of dispositional confidence and ultimately the traits thought to underpin MT.

6.4 General Conclusion

The aim of Study 4 was to further examine the within- and between-network properties of the HCQ-R by examining its test-retest reliability, convergent validity, and predictive validity. Although Investigation 1 did not provide support for the HCQ-R's convergent validity when using MacKenzie *et al.'s* (2011) and Marsh's (2007) conservative guidelines, the findings did satisfy Kline's (2005) liberal thresholds and were comparable to parameters used in recent research. Consequently, the findings of Study 4 appear to provide preliminary evidence to support the HCQ-R's convergent validity. In contrast, the findings from Investigation 1 revealed little support for the HCQ-R's test-retest reliability and no support for its predictive validity when assessed with coping styles, which in turn, casts doubt over the instrument's stability and its ability to predict constructs within its nomological network. Possible explanations for these findings are discussed. However, Investigation 2 provided preliminary evidence to support the HCQ-R's predictive validity whereby test scores significantly predicted pre-competitive challenge states and pre-competitive self-confidence and somatic anxiety. Given the importance placed on challenge states, self-confidence, and anxiety management for the mentally tough performer (Bull *et al.*, 2005; Coulter *et al.*, 2010; Fourie & Potgieter, 2001; Gucciardi *et al.*, 2008; Jones *et al.*, 2002, 2007; Thelwell *et al.*, 2005), the findings of Investigation 2 suggest that the traits thought to underpin MT may provide meaningful information to predict salient pre-competitive MT characteristics. However, given the discussed impact of situational factors in predicting cognitions and behaviours, caution should be exercised when interpreting the extent to which HCQ-R scores predict pre-competitive cognitions. One possible limitation to Study 4 is the relatively small sample sizes used to examine the HCQ-R's predictive validity. Given that the samples used only marginally satisfied Green's (1991) criteria, the generalisability of the regression models may be limited (Tabachnick & Fidell, 2001). Research using larger (e.g., n = 200-300) samples is required to enhance our understanding of the HCQ-R's predictive validity.

CHAPTER 7.0

The primary aim of this thesis was to examine the factorial validity of the MTQ48 in an effort to provide a valid measure of MT. Despite the MTQ48 being the most utilised measure of MT in sport, preliminary evidence regarding its factorial validity has been equivocal (Gucciardi *et al.,* 2012; Horsburgh *et al.,* 2009. This has resulted in some researchers (e.g., Connaughton & Hanton, 2009; Gucciardi *et al.,* 2011) questioning the adequacy of the MTQ48 to measure the 4/6C's model of MT (Clough *et al.,* 2002; Earle, 2006). In light of these equivocal findings, researchers may be prematurely dismissing the MTQ48 as a valid measure of the 4/6C's model of MT, and in turn, may be prematurely dismissing the 4/6C's model of MT as a valid conceptualisation of MT. The present thesis therefore sought to provide a comprehensive examination of the factorial validity of the MTQ48 in an effort to provide an adequate measure of the 4/6C's model of MT. The 4/6C's model of MT provided a theoretical underpinning for the thesis, and the programme of studies examined its utility in providing a means to assess the traits which may underpin MT.

Additionally, the present thesis sought to replicate and extend the research of Gucciardi *et al.* (2012) by using a larger sample of athletes to examine the factorial validity of its various underpinning factor structures. The present thesis makes three unique contributions to the MT literature. First, Study 1 provided a more comprehensive examination of the MTQ48's factorial validity in sport. Second, Study 2 developed an instrument to better represent the 4/6C's model of MT and provided an examination of its factorial validity. Third, Studies 3-4 developed an instrument to better represent the core traits underpinning the 4/6C's model of MT and provided an examination of its construct validity. The aim of this chapter is to provide a summary of the findings emerging from these studies, highlight the implications of the research, and identify avenues for future research.

7.1 Summary of Findings

The first specific aim of this thesis was to examine the factorial validity of the first- and second-order and four and six factor models of the MTQ48. Study 1 found little support for the factorial validity of the MTQ48. The data gleaned from Study 1 adds to the research by Gucciardi *et al.* (2012) which suggests that the first-order four factor model structure of the MTQ48 is not a valid measure of the 4/6C's model of MT which it intends to capture. In addition, Study 1 extended the research by Gucciardi *et al.* and found little support for its alternative model structures (six factor model, first- and second-order representations). Consistent with Gucciardi and colleagues, multiple criteria (factor loadings, multiple model fit indices, face validity) were used to evaluate the factorial validity of the MTQ48. According to the International Test Commission (ITC; 1999), users should "use tests only for those purposes which relevant and appropriate validity evidence is available" (p.12). Indeed, if no information is available to support an instrument's factorial validity, its validation cannot be progressed and users of the instrument cannot have confidence in interpreting its scores (Gignac, 2009). This has resulted in some researchers (e.g., Connaughton & Hanton, 2009; Gucciardi *et al.*, 2011) questioning the adequacy of the MTQ48 to measure the 4/6C's model of MT (Clough *et al.*, 2002; Earle, 2006).

On the proviso that the MTQ48 may exhibit poor factorial validity, the second specific aim of this thesis was to provide a valid tool to measure the 4/6C's model of MT. In Study 2, the UCMTQ was developed whereby items better represented the definitions forwarded by Clough *et al.* (2007) to underpin the 4/6C's model of MT. The findings of Study 2 demonstrate that the use of rigorous item generation procedures (e.g., Clark & Watson, 1995; MacKenzie *et al.*, 2011) aimed at adequately capturing Clough *et al.'s* definitions resulted in superior model fit when compared to the tested models of the MTQ48. However, despite UCMTQ items conforming to recommendations in the item generation literature and item content being supported by six independent scrutineers, findings revealed little support for the UCMTQ's factorial validity. Given that psychometric tools serve to validate its underpinning conceptualisation (Marsh, 2002), this finding indicates that the 4/6C's model of MT requires revision and at present, does not provide a valid conceptualisation of the traits thought to underpin MT. Study 2 concluded that the UCMTQ's poor factorial validity was attributable to poorly constructed definitions to underpin the 4/6C's model of MT. Scale developers have emphasised the importance of adequately defining the construct and its conceptual domain because:

"…many researchers *think* they have a clear idea of what they wish to measure, only to find out that their ideas are more vague than they thought. Frequently, this realization occurs after considerable effort has been invested in generating items and collecting data—a time when changes are far more costly than if discovered at the outset of the process" (DeVellis, 1991, p. 51).

Devellis' (1991) commentary appears to have profound links to the evolution of the MTQ48 and the 4/6C's model of MT whereby problems have been identified with the instrument's conceptual make-up long after its conception and usage in the literature. In light of the documented concerns regarding the MTQ48's conceptual make-up (e.g., Connaughton & Hanton, 2009; Gucciardi *et al.*, 2011) and the lack of support for its factorial validity, Gucciardi *et al.* (2012) suggested that the entire framework of the instrument needs to be reconsidered. Although concerns regarding the MTQ48's factorial validity are substantiated, the dismissal of the entire trait conceptualisation underpinning the 4/6C's model of MT appears to be somewhat premature given that Gucciardi *et al.* only tested the factorial validity of the MTQ48 items as opposed to examining the core traits underpinning the instrument.

In Study 3, the utility of the trait conceptualisation underpinning the 4/6C's model of MT was examined. A key finding of Study 3 was that preliminary support was found for the factorial validity and dimensionality of an instrument designed to measure the core traits underpinning Clough *et al.'s* (2002) 4C's model of MT. The findings of Study 3 illustrate that the use of rigorous scale development procedures (e.g., Clark & Watson, 1995; Hinkin, 1995; MacKenzie *et al.*, 2011; Sartori, 1984) aimed at adequately defining dispositional challenge, commitment, control, and confidence resulted in superior model fit when compared to the respective models of the MTQ48 and UCMTQ. Specifically, findings showed acceptable levels of factorial validity of the first-order four factor revised model of the HCQ-R. However, it is important to note that factorial validity was only acceptable when using the most liberal model fit thresholds, therefore providing preliminary support for the factorial validity of the HCQ-R. Notwithstanding, given that the HCQ-R is underpinned by hardiness and more liberal thresholds have been used to support the factorial validity of the DRS15-R (Hystad *et al.*, 2010), the findings of Study 3 suggest that the model fit of the HCQ-R is comparable to existing measures of hardiness. The findings of Study 3 therefore emphasise the importance of adequately defining the target construct prior to developing an operational tool (e.g., Devillis, 1991; Gignac, 2009) and provide preliminary support for the within-network properties of an instrument designed to measure the traits thought to underpin MT.

Consistent with the construct validity enterprise (Marsh, 2002), the findings of Study 3 provided a platform and pre-requisite for the examination of the between-network properties of the HCQ-R (Gignac, 2009; Marsh & Hau, 2007). A key finding of Study 4 was demonstrating preliminary support for the HCQ-R's predictive validity. According to Andersen *et al.* (2007), predictive validity provides the most meaningful information in understanding the scores yielded from psychometric tools. Although findings did not support the HCQ-R's ability to predict coping styles in athletes, support was found for its ability to predict pre-competitive challenge states and pre-competitive self-confidence and somatic anxiety. Study 4 therefore provides preliminary support for the HCQ-R's ability to predict established MT characteristics (Bull *et al.*, 2005; Coulter *et al.*, 2010; Fourie & Potgieter, 2001; Gucciardi *et al.*, 2008; Jones *et al.*, 2002, 2007; Thelwell *et al.*, 2005). Thus, the findings of Studies 3-4 not only indicate preliminary support for the factorial validity and dimensionality of the HCQ-R, but also suggest that the instrument may have some promise regarding its practical utility. Findings also provide some support for the HCQ-R's convergent validity, thus enhancing our confidence that the measure appears to capture what it intends to capture (MacKenzie *et al.*, 2011). However, the convergent validities obtained were only acceptable when using more liberal thresholds (Kline, 2005) and correlation patterns which highlights the need to exercise some caution when interpreting these findings.

Although Studies 3-4 provide preliminary support for the HCQ-R's validity when using liberal thresholds, the findings relating to the instrument's reliability were less supportive. Specifically, Study 4 provided little evidence to support the test-retest reliability of the HCQ-R, thus questioning the temporal stability of the instrument. Response variation was offered as a possible explanation for the lack of test-retest reliability, whereby participants could have been unclear about how they felt when responding to HCQ-R items across two time points (Nevill *et al.*, 2001). Despite the importance placed on establishing test-retest reliability in personality scales (e.g., McCrae *et al.*, 2011; Nevill *et al*; Watson, 2004), there is limited discussion in the literature regarding possible explanations as to why an instrument may display poor stability. Indeed, McCrae *et al.* posed a number of key questions which need to be considered when assessing test-retest reliability of personality inventories. Why do respondents choose different answers on different occasions? Why do they do so more with some traits than others? According to McCrae and colleagues, these questions are central to our understanding of personality assessment that have not yet been addressed and require future research attention.

A possible reason to explain the equivocal construct validity findings of the HCQ-R could have been due to the environment at which the questionnaires were completed. One potential limitation of Studies 1-3 was that data collection procedures were invariably conducted with large groups of participants (50 – 300) at the beginning or end of University lectures and seminar sessions. Research from the applied setting has suggested that athletes will be more willing to complete questionnaires when they are in an undisturbed and relaxing environment which affords no time restrictions (Woodcock, Duda, Cumming, Sharp, & Holland, 2012). Despite there being a fully briefed researcher present, the large group sizes may have prevented participants from asking clarifying questions in fear of being negatively evaluated by a large number of their peers. Moreover, given that the questionnaires were completed within University lectures and seminar sessions, students may have felt as if there was a time restriction which could have limited their attention to the questionnaires. Consequently, it is possible that group size and the test environment could have compromised response validity in Studies 1-3. In a similar vein, response validity could have been compromised in Study 4. Although participants completed the questionnaires in small groups (i.e., with their sport teams), participants may not have been willing to ask clarifying questions regarding their pre-competition cognitions in the fear of being negatively evaluated by their teammates. Furthermore, given that participants were asked to complete the questionnaires prior to training and competition in Investigation 2, participants may have exercised undue care and consideration in responding to questions as their attention may have been focused on preparing for the upcoming competition. Despite the inherent need to collect very large amounts of data, scale developers are encouraged to consider the location of test administration and use small groups (e.g., 20-30) to promote maximal respondent engagement and ultimately facilitate high response validity.

Construct validation is fundamental to the development and validation of psychometric tools, which ultimately serves to validate its underpinning conceptualisation (Marsh, 2002). Although the findings of Study 3-4 provide preliminary support for the HCQ-R's construct validity when using more liberal thresholds, the equivocal predictive validity and test-retest reliability findings cast doubt regarding the adequacy of the instrument to measure the traits thought to underpin MT. Given that construct validation is an ongoing examination of both within- and between-network properties beyond those methods employed in Studies 3-4 of this thesis, further examination of the HCQ-R's construct validity is therefore needed before the instrument can be considered a useful tool to measure MT. Specific attention is required to examine the HCQ-R's test-retest reliability and predictive validity, with greater emphasis being placed on its ability to predict behavioural measures of MT and performance (Andersen *et al.*, 2007).

7.2 Implications

Mental toughness and hardiness

This program of research has a number of implications relating to the MT and hardiness literature. The findings of Study 1 add to the concerns regarding the practices used to develop the MTQ48 and indicate little support for its utility as a valid measure of the 4/6C's model of MT. Although Study 2 revealed marginally superior model fit for the six factor models when compared to the four factor models of MT, little support was found for the factorial validity of an instrument designed to measure the 4/6C's model of MT. Despite the MTQ48 being the most utilised measure of MT in sport (Gucciardi *et al.*, 2011), the evidence gleaned in Studies 1-2 indicates that the instrument and its underpinning conceptualisation are not valid and ultimately, not fit for purpose. The development of the revised 4C's model of MT provides a promising platform to examine the core traits thought to underpin MT. However, the findings of Studies 3-4 only provide preliminary support for its utility. Consequently, the HCQ-R is not ready to be used as an operational tool (e.g., talent identification/selection) for research and consultancy. In line with Marsh (2002), scale development is an ongoing process and more research is needed to examine the HCQ-R's utility as an instrument to measure MT in sport. Although further examination of the HCQ-R in the areas previously discussed (i.e., test-retest reliability, predictive validity) may enhance our understanding of the validity and reliability of the instrument, it is important to note that it could also result in no discernible improvements to the evidence gleaned in Studies 3-4.

One important consideration in progressing the understanding of the revised 4C's model of MT concerns its conceptual make-up. Despite being embedded in the 3C's of hardiness, the strong conceptual relations among dispositional challenge, commitment, and control could have contributed toward the equivocal construct validity findings of the HCQ-R. According to Kobasa *et al.* (1982), challenge is characterised by a tendency to believe that change is normal in life and that change provides opportunities to develop rather than threats to security. Commitment is reflected by a tendency to feel involved in whatever one is doing and encounters (Kobasa *et al.*). Control is characterised by a tendency to be influential when faced with difficulties (Kobasa *et al.*). When the relations among the respective hardiness subcomponents are analysed, it is theoretically plausible to assume substantial overlap in conceptual space. For instance, if a person tends to feel influential when faced with difficulties (control), one would predict that person to generally anticipate changes in their life as manageable and less threatening (challenge). Moreover, if a person tends to feel deeply involved in the activities they undertake (commitment), one would predict that person to generally feel influential over their environment (control) and generally feel that they can manage unexpected difficulties (challenge). These examples demonstrate the close theoretical relations among the 3C's of hardiness. Consequently, although Kobasa (1979) theorised hardiness to be underpinned by independent yet related subcomponents, one may tentatively suggest that the level of conceptual independence may be weaker than first hypothesised.

Indeed, Hystad *et al.* (2010) only found support for the factorial validity of the DRS15-R when using more liberal model fit evaluation criteria. Given that an extremely large (n = 7,280) and homogenous sample of Norwegian Armed Forces was used, one could expect more convincing within-network properties of the DRS15-R. Hystad *et al.'s* factorial validity evidence appears to provide only preliminary support for the factor structure of hardiness which suggests that the strong conceptual relations among dispositional challenge, commitment, and control may compromise not only the factor structure of the 3C's of hardiness, but also the revised 4C's model of MT. Given this proposition, it could be argued that the addition of dispositional confidence further compounds issues relating to the conceptual coverage of the revised 4C's model of MT. Although Clough and colleagues' inclusion of dispositional confidence in a model to underpin MT in sport is intuitively compelling, it is theoretically plausible for confidence to be closely related to all three hardiness subcomponents (see Chapter 2.4.2.9 for a comprehensive discussion on the relations among hardiness and confidence). Consequently, a potential limitation of the revised 4C's model of MT could be that its underlying subcomponents capture a vast expanse of conceptual space in the literature. The overlapping nature of the constructs in question may promote blurred conceptual boundaries between factors which may ultimately explain its compromised factor structure and construct validity.

Although many sport psychology practitioners believe that there is a construct (made-up of traits and/or a combination of characteristics) that constitutes MT, some researchers have raised concerns regarding the overarching nature of MT (e.g., Andersen, 2011). These concerns stem from the overwhelmingly large number of attributes identified in the literature to make-up MT. Indeed, qualitative studies (Bull *et al.*, 2005; Coulter *et al.*, 2010; Gucciardi *et al.*, 2008; Jones *et al.*, 2002, 2007; Thelwell *et al.*, 2005) have identified 118 attributes thought to make-up the mentally tough performer. Specifically, Andersen argues that MT appears to encapsulate repackaged constructs which on the surface appear novel, yet are actually extensively discussed in the literature. According to Andersen, the proliferation of MT research appears to stem from its positive appeal to key stakeholders in sport (e.g., media, athletes, coaches, sport psychologists). The issue of established constructs being repackaged to represent novel constructs has received ample discussion in the psychology literature (e.g., Block, 1996, 2000). It appears that MT has suffered from both jingle (Thorndike, 1904) and jangle (Kelley, 1927) fallacies. The jingle fallacy refers to the use of common terms for different constructs whereas the jangle fallacy refers to the use of different terms for common underlying constructs. Specifically, the jangle fallacy appears to be particularly prevalent in the dispositional school of thought of MT whereby hardiness has been used to essentially examine a construct equivalent to MT. Given the criticisms targeted at MT, it is clear that the area suffers from a lack of conceptual clarity. Although the rigour underpinning the examination of MT has led to more robust findings, the diverse perspectives of MT pose a significant barrier to the advancement of knowledge in the area. At present, researchers appear to be more concerned with examining a particular model of MT as opposed to advancing a shared understanding of MT. According to Gucciardi and Gordon (2011), the formulation of a common definition of MT may facilitate reconciliation of perspectives and serve to unite a fragmented area of research. However, given the ongoing debates in the academic community regarding the problematic overarching nature of MT, one of the most significant challenges for this area is to convince the academic community of its empirical and practical value.

Scale development and validation

The present thesis may also have important implications in extending our knowledge base of best practices regarding scale development and validation within sport and exercise psychology. Although many key scale development and validation articles have been harnessed from other related academic domains (e.g., Educational Psychology: Marsh, 1997, 2002, 2007), there appears to be little up-to-date information reported in the sport and exercise psychology literature on how best to develop a psychometrically sound questionnaire. Given the need to develop valid and reliable psychometric instruments to validate theory both in the MT (e.g., Gucciardi *et al.*, 2011) and in the sport and exercise psychology literature (e.g., Hagger & Chatzisarantis, 2009; Marsh, 2002), one could argue that the literature base could benefit from such information. Adopted from the marketing literature, MacKenzie *et al.'s* (2011) scale development framework provided a comprehensive step-by-step guide to underpin the development and evaluation of the respective questionnaires within this thesis. The framework provides an amalgamation of the most up-to-date best practices in scale development and validation. Consequently, MacKenzie *et al.'s* work provides researchers with a promising framework to guide the development and validation of psychometrically sound tools in sport and exercise psychology. This thesis highlights the need for future research to use sound procedures when developing and validating psychometric tools as failure to do so can have detrimental consequences to theory development and ultimately the advancement of knowledge. It is therefore imperative that researchers harness information from multiple domains to facilitate best practice in the development and validation of psychometric tools, which in turn, will ultimately facilitate the central objective of research – knowledge advancement.

Another important implication from this thesis relates to best practice of the parameters used to guide the construct validity process. Although the existing literature provides information to guide the validation process (e.g., Marsh, 1997, 2002, 2007), the various thresholds outlined makes it difficult to accurately evaluate the adequacy of any given instrument. From this thesis, there are two key components of the validation process which need to be addressed; factorial validity and convergent validity. With regards to factorial validity, there is an ongoing debate in the literature regarding the use of conservative and liberal thresholds to guide model fit evaluation. Despite some researchers advocating the use of more conservative parameters (e.g., Hu & Bentler, 1999; Russell, 2002), others have emphasised the need to use more liberal guidelines (e.g., Marsh *et al.,* 2004). The use of various model fit parameters has been reflected in the hardiness and MT literature whereby researchers have used more liberal parameters (e.g., Hystad *et al.,* 2010) and both conservative and liberal thresholds (e.g., Gucciardi *et al.,* 2009; Gucciardi & Gordon, 2009; Gucciardi *et al.,* 2012) to guide model fit evaluation. Furthermore, research has failed to state the parameters by which model fit evaluation has been assessed (Sheard *et al.,* 2009). Research in sport and exercise psychology has also used different means to assess model fit including the use of conservative thresholds (e.g., Bartholomew *et al.,* 2010; Freeman *et al.,* 2011; Williams & Cumming, 2011), both liberal and conservative thresholds (e.g., Coffee & Rees, 2008; Lonsdale *et al.,* 2008; Marsh *et al.,* 2010) and a failure to specify (e.g., Zourbanos *et al.,* 2009). Typically, conservative thresholds are indicative of good model fit and liberal thresholds are indicative of adequate or acceptable model fit. Despite these apparent classifications, the interchangeable use of phraseology invariably makes it difficult for the reader to make accurate judgements about an instrument’s factorial validity. Consequently, the use of multiple model fit evaluation guidelines may not only call into the question the rigour underpinning the development and validation of psychometric instruments, but also the conclusion gleaned from the data collected. Given the importance placed on establishing an instrument's factorial validity (Gignac, 2009; Marsh & Hau, 2007), consensus is needed regarding the use of conservative and liberal thresholds when evaluating model fit to enhance our understanding of psychometric instruments and ultimately the utility of theory.

In a similar vein, there are also multiple parameters outlined in the literature to guide the evaluation of convergent validity. According to MacKenzie *et al.* (2011) and Marsh (2007), strong relationships (r = 0.70) are indicative of convergent validity, whereas Kline (2005) suggests using more liberal thresholds (r = 0.30 to 0.50). However, it is important to highlight that subtle differences are present in the definitions used to describe convergent validity between these recommendations. Specifically, MacKenzie *et al.* and Marsh state that convergent validity involves correlating instruments which purport to measure the same or substantially overlapping constructs to the target instrument. Thus, from a theoretical perspective, it is intuitive to expect strong correlations (i.e., $r > 0.70$). However, Kline states that nomological validity is indicated by weak to moderate correlations. Given that nomological validity concerns the assessment of the lawful relationships within a construct's theoretical framework (Cronbach & Meehl, 1955), it does not necessarily involve constructs which directly overlap with the target construct. Therefore, one may also argue that Kline's more liberal threshold is theoretically justified.

In light of these recommendations, one would expect researchers to use the aforementioned criteria according to whether the aim of the study is to examine convergent validity or nomological validity. However, given the blanket use of convergent validity to represent these two types of validity, research in sport and exercise psychology has generally used more liberal criteria to indicate support for convergent validity irrespective of the theoretical associations between the constructs in question. Specifically, many studies examining MT (e.g., Gucciardi *et al.*, 2009; Gucciardi & Gordon, 2009) and beyond (e.g., Freeman *et al.*, 2011; Lonsdale *et al.*, 2008; Williams & Cumming, 2011; Zourbanos *et al.*, 2009) have used correlation patterns to indicate support for convergent validity. Although the use of correlation patterns appears to provide an intuitive means to evaluate convergent validity, this method is open to criticism as there is no clear procedure to underpin the acceptance or rejection of study hypotheses. At present, there is a clear distinction between the parameters outlined in the scale validation literature and the methods being used in sport and exercise psychology research. Consequently, the rigour underpinning the assessment of convergent validity in the sport and exercise psychology literature is questionable and appears to compromise the meaning of the conclusions gleaned.

In order to enhance our understanding of the construct validation process, consensus is needed regarding best practices in the procedures used to evaluate factorial validity and convergent validity. Greater clarity regarding best practices in the construct validity process would give researchers and practitioners more confidence in the instruments they develop and adopt, which in turn, may enable more robust conclusions to be drawn from psychometric data. This is paramount given that psychometric instruments are often the central means for theory testing and development (Marsh, 2002). This could also provide a clearer means to assess and monitor the development of instruments/theory over time. This is especially important as construct validation is an ongoing process (Marsh, 2002) whereby instruments are often adapted and revised in light of empirical data (e.g., theory development) and their practical use (e.g., completion time). It is therefore apparent that more research is needed to strengthen the guidelines used to develop psychometric tools and the theory it serves to validate.

7.3 Limitations and Recommendations for Future Research

The present thesis makes a unique contribution to the MT literature by a) providing a comprehensive examination of the MTQ48's factorial validity, and b) providing the development and validation of an instrument to measure the traits thought to underpin MT. The aim of this section is to outline the limitations of the thesis and inform recommendations for future research. The strengths of this program of research include the use of MacKenzie *et al.'s* (2011) scale development framework, the relatively large sample sizes used (especially Study 1), the comprehensive examination of the MTQ48's factorial validity, and the examination of both within- and between-network properties of an instrument designed to measure the traits thought to underpin MT.

Given that construct validation is an ongoing examination of both within- and between-network properties (Marsh, 2002), research is required to further examine the psychometric properties of the HCQ-R. One limitation of the HCQ-R was that preliminary evidence did not support its test-retest reliability. In accordance with Watson (2004), researchers are encouraged to assess the temporal stability of instruments over multiple time points (e.g., 1, 3, and 6 weeks) to facilitate a more thorough examination of test-retest reliability. Given that clarity of item responses was offered as a reason to explain the HCQ-R's poor test-retest reliability, think aloud protocols (Dietrich & Ethrenspiel, 2010) whereby participants verbally report their thoughts as they respond to HCQ-R items could provide an insightful means to understand the reasons underpinning poor temporal stability. For instance, comparing think aloud responses over multiple time points may offer a fruitful means to understand why participants choose different answers on different occasions (McCrae *et al.,* 2011). Despite the importance placed on examining test-retest reliability (especially when examining personality traits), it is often overlooked in the construct validation process due to the need to collect data over multiple time points (e.g., McCrae *et al*; Nevill *et al*., 2001; Watson). The inconsistent assessment of test-retest reliability is reflected in the MT literature whereby only limited research has examined stability (Earle, 2006), while others have used internal consistency as a means to assess reliability (e.g., Gucciardi *et al.*, 2009; Gucciardi & Gordon, 2009). In order to enhance the integrity of the conclusions drawn from questionnaires, researchers in MT and beyond are strongly encouraged to incorporate test-retest reliability as routine practice in the development and validation of instruments.

Although Study 4 provides preliminary evidence to support the HCQ-R's ability to predict pre-competitive cognitions, no information exists regarding its relationship with behaviour and performance. According to Andersen *et al.* (2007), establishing the link between test scores and actual real-world behaviours is essential to the understanding and application of data. That is, unless test scores are calibrated in some way against overt behavioural variables, the meaning of those test scores is unclear. In order to truly demonstrate the HCQ-R's utility, research is required to examine its ability to predict both behavioural measures of MT and performance. Future research using experimental designs could examine the relative main and interactive effects of HCQ-R scores upon perseverance time and performance of a motor task. Specifically, participants could be asked to complete a putting task on two occasions whereby an unachievable target score is set. Putting accuracy could be measured on both occasions and task perseverance (mentally tough behaviour) could be monitored by assessing their willingness to participate in optional free practice sessions prior to the second testing session. Additionally, the effects of HCQ-R scores on perseverance and endurance time could be examined. Participants could be asked to complete an endurance task whereby an unachievable target is set relative to their aerobic capacity. Task perseverance could be measured by their willingness to continue to reach their assigned goal (i.e., endurance time). Researchers are strongly encouraged to examine the predictive validity of self-report tests relative to actual behaviour and performance to enhance future validation procedures (Andersen *et al.*, 2007) and ultimately the utility of theory to the sporting context.

Research is also required to examine the HCQ-R by using methods beyond the scope of this thesis. Measurement invariance is another important statistical and methodological consideration when evaluating the psychometric properties of the HCQ-R (Cheung & Rensvold, 2002). Specifically, measurement invariance has important implications for making valid comparisons of group means. If the factor structure of the HCQ-R is invariant across groups, one will have more confidence that comparisons between groups accurately reflect real group differences, which in turn, may provide a promising platform to examine group differences in MT traits. Given that inferences cannot be made about invariant measurement properties for subgroups of a population (Horn & McArdle, 1992), research is required to examine whether the HCQ-R's factor structure is invariant across various athlete demographics including age, gender, skill level, experience, and sport. Indeed, research by Golby and Sheard (2004) found that International level rugby league players scored significantly higher on all three hardiness subcomponents when compared to Division One players. Statistical techniques such as multi-sample CFA will provide a more stringent examination of the strength and generalisability of the revised 4C's model of MT. Researchers are encouraged to use complex statistical techniques such as multi-sample CFA to strengthen the validation procedures used in future sport and exercise psychology research.

A further recommendation for future scale development and validation research is the need to recognise the factors which influence response validity. Social desirability has been argued to be one of the most prominent sources of systematic error which may compromise response validity (Podsakoff *et al.,* 2003). Despite there being little impact of social desirable responding in this thesis, researchers are encouraged to examine the impact of this variable on item responses and ascertain the likelihood of participants to respond in an honest manner. In line with the procedures used in this thesis, test administrators should emphasise that there are no correct or incorrect answers to the questionnaire and clearly state that all responses are anonymous to reduce the likelihood of participants feeling the need to respond in a certain manner. Despite a number of recent studies examining social desirability in the development of questionnaires in sport and exercise psychology (e.g., Freeman *et al.,* 2011; Gucciardi *et al.,* 2009), this area has generally been overlooked and requires more attention in future research.

In a similar vein, it is important to highlight the potential impact of delivery methods upon response validity. Recent research in sport and exercise psychology has used a variety of delivery methods including supervised paper (e.g., Freeman *et al.,* 2009; Williams & Cumming, 2011), online (e.g., Gucciardi & Gordon, 2009; Gucciardi *et al.,* 2012), and a combination of supervised paper and postal (e.g., Bartholomew *et al.,* 2010) and online and postal (e.g., Lonsdale *et al.,* 2008). Although the administration of online questionnaires provides a promising remedial solution against the adverse impact of time constraints and test environment on response validity, researchers need to recognise the importance of having a fully briefed test administrator available to clarify potential problems with test completion (Cooper, 2010; Kline, 2005). According to the ITC's (1999) International Guidelines for Test Use, test administrators should know "how to deal with a test taker's questions during test administration" (p. 9) and "administer tests under appropriate supervised conditions" (p. 14). Indeed, the inherent need to collect and input very large amounts of data places significant demands on the researcher. However, supervised delivery methods enable the researcher to have more control over respondents' test-taking behaviour, which should be considered more important than the demands placed on the test administrator. Research is therefore strongly encouraged to use supervised delivery methods when developing instruments in an effort to enhance response validity and ultimately the conclusions gleaned from questionnaires.

One key consideration highlighted in this program of research is the importance of adequately defining the target construct(s) prior to engaging in scale development and validation. The scale development literature (e.g., Clark & Watson, 1995; Hinkin, 1995; MacKenzie *et al.*, 2011; Sartori, 1984) emphasises the need for researchers to fully understand the construct in question, as failure to adequately define a construct undermines construct validity, internal validity, and statistical conclusion validity (MacKenzie, 2003). Indeed, "To move too quickly to potentially superficial between-construct research is to risk within-construct problems that characterise many psychological measures" (Marsh *et al.*, 2010, p. 464). The inherent problems associated with inadequately defining a conceptual model are reflected in the findings of this thesis. Clough and colleagues appear to have compromised the conceptual representation of the 4/6C's model of MT by moving too quickly in developing an operational instrument to capture the traits thought to underpin MT. Although it is important to note that MacKenzie *et al.'s* framework was not available to Earle (2006) when developing the MTQ48, the literature encompassed within this framework (e.g., Churchill, 1979; Clark & Watson; Hinkin; Kline, 2000; MacKenzie; Sartori) was available. Given the importance placed on adequately defining latent constructs, it is imperative that future research devotes more time and resources in understanding the construct(s) in question prior to its psychometric examination. Similar to the processes adopted in this thesis, research is strongly encouraged to conduct comprehensive literature searches to better understand the target construct and how it relates to other established constructs (Clark & Watson). This in turn may enable researchers to better understand the target construct's conceptual underpinning and ultimately enhance the likelihood of developing theoretically embedded instruments.

7.4 Conclusions

In summary, the findings from Study 1 of this thesis indicate that the MTQ48 is not a valid measure of MT in sport. The findings are consistent with the research by Gucciardi *et al.* (2012), and this thesis has extended this research by providing a more comprehensive examination of the MTQ48's factorial validity. The findings from Study 2 provide evidence to indicate that the MTQ48's underpinning factor definitions (Clough *et al.,* 2007) are conceptually insufficient. The collective findings of Studies 1-2 indicate that the MTQ48 and its underpinning 4/6C's model of MT require revision and at present, do not provide valid means to examine the traits thought to underpin MT in sport. Although Studies 3-4 provide preliminary support for the HCQ-R's validity, the findings were less convincing for the instruments reliability. Consistent with the construct validation approach (Marsh, 2002), the equivocal findings of Studies 3-4 indicate that further understanding of the HCQ-R's construct validity is warranted before it can be considered a useful tool to measure MT in sport. Despite the efforts of the thesis, considerable work is needed before we can have confidence in a trait measure of MT. In order to achieve this, it is imperative that future research uses the scale development literature to guide this journey and ultimately facilitate knowledge advancement in the field of MT in sport.

REFERENCES

Abbott, A., & Collins, D. (2004). Eliminating the dichotomy between theory and practice in talent identification and development: considering the role of psychology. *Journal of Sports Sciences*, **22**, 395-408.

Adie, J.W., Duda, J.L., & Ntoumanis, N. (2008). Achievement goals, competition appraisals, and the psychological and emotional welfare of sport participants. *Journal of Sport & Exercise Psychology*, **30**, 302–322.

Aidman, E.V. (2007). Attribute-based selection for success: The role of personality attributes in long term predictions of achievement in sport. *The Journal of the American Board of Sport Psychology*, **3**, 1-18.

Aidman, E. V., & Schofield, G. (2004). Personality and individual differences in sport. In T. Morris & J. Summers (Eds.), *Personality and individual differences in sport* (pp. 22-47). Milton: Wiley.

Aish, A. M., & Jöreskog, K. G. (1990). A panel model for political efficacy and responsiveness: An application of LISREL 7 with weighted least squares. *Quality and Quantity,* **24,** 405-426.

Akaike, H. (1987). Factor analysis and AIC. *Psychometrika,* **52,** 317-332.

Alderman, R. B. (1974). *Psychological behaviour in sport.* Toronto: W.B. Saunders Company.

Allen, M. S., Greenlees, I., & Jones, M. (2011). An investigation of the five-factor model of personality and coping behaviour in sport. *Journal of Sports Sciences,* **8,** 841-850.

Anastasi, A., & Urbina, S. (1997). *Psychological Testing* (7th ed.). New York: Prentice-Hall.

Andersen, M. B. (2011). Who’s mental, who’s tough, and who’s both?: Mutton constructs dressed up as lamb. In S. Gordon & D. J. Gucciardi (Eds.), *Mental toughness in sport: Developments in research and theory* (pp. 66-85). London: Routledge.

Andersen, M. B., McCullagh, P., & Wilson, G. J. (2007). But what do the numbers really tell us?: Arbitrary metrics and effect size reporting in sport psychology research. *Journal of Sport & Exercise Psychology,* **29,** 664-672.

Antonovsky, A. (1974). Conceptual and methodological problems in the study of resistance resources and stressful life events. In B. Dohrenwend & B. Dohrenwend (Eds.), *Life Events: Their Nature and Effects* (pp. 245-258). New York: Wiley.

Arbuckle, J. L. (2009). *Amos 18 user’s guide*. Crawfordville, FL: Amos Development Corporation.

Averill, J. R. (1973). Personal control over aversive stimuli and its relationship to stress. *Psychological Bulletin,* **80,** 286-303.

Bandalos, D. L., & Boehm-Kaufman, M. R. (2008). Four common misconceptions in exploratory factor analysis. In Lance, C. E., & Vandenberg, R. J. (Eds.), *Statistical and Methodological Myths and Urban Legends: Doctrine, Verity and Fable in the Organizational and Social Sciences* (pp. 61–87). Taylor & Francis.

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

Bandura, A. (1997). *Self-efficacy: The exercise of control.* New York: W. H. Freeman and company.

Bartholomew, K. J., Ntoumanis, N., & Thøgersen-Ntoumani, C. (2010). The controlling interpersonal style in a coaching context: development and initial validation of a psychometric scale. *Journal of Sport and Exercise Psychology,* **32**, 193-216.

Bartone, P. T. (1995, July). A short hardiness scale. Paper presented at the 103rd Annual Convention of the American Psychological Association, New York, NY.

Bartone, P. T. (1999). Hardiness protects against war-related stress in army reserve forces. *Consulting Psychology Journal,* **51**, 72-82.

Bartone, P. T., Ursano, R. J., Wright, K. W., & Ingraham, L. H. (1989). The impact of a military air disaster on the health of assistance workers: A prospective study. *Journal of Nervous and Mental Disease,* **177**, 317–328.

Bartram, D. (1994). *PREVUE assessment technical manual* (2nd ed.). Vancouver, BC: Prevue Assessments International.

Beauducel, A., & Herzberg, P. Y. (2006). On the performance of maximum likelihood versus means and variance adjusted weighted least squares estimation in CFA. *Structural Equation Modeling,* **13**, 186–203.

Beedie, C.J., Terry, P.C., & Lane, L.M. (2000). The profile of mood states and athletic performance: Two eta-analyses. *Journal of Applied Sport Psychology,* **12**, 49-68.

Bentler, P.M. (1990). Comparative Fit Indexes in Structural Models. *Psychological Bulletin*, **107**, 238-246.

Bentler, P. M. (1992). On the fit of models to covariances and methodology to the Bulletin. *Psychological Bulletin*, **112**, 400-404.

Bentler, P. M. (1995). EQS structural equations program manual. Encino, CA: Multivariate software.

Bentler, P. M. (2006). EQS 6 structural equations program manual. Encino, CA: Multivariate Software.

Berlyne, D. E. (1964). Novelty. *New Society,* **87,** 23-24.

Birnbaum, M. H. (2004). Human research and data collection via the Internet. *Annual Review of Psychology*, **55**, 803–32.

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

Blascovich, J., Seery, M. D., Mugridge, C. A., Norris, R. K., & Weisbuch, M. (2004). Predicting athletic performance from cardiovascular indexes of challenge and threat. *Journal of Experimental Social Psychology,* **40**, 683–688.

Block, J. (1996). Some jangly remarks on Baumeister and Heatherton. *Psychological Inquiry*, **7**, 28-32.

Block, J. (2000). Three tasks for personality. In L. R. Bergman, R. B. Cairns, L. G. Nilsson & L. Nystedt (Eds.), *Developmental Science and the Hollistic Approach* (pp. 155-164). Mahwah, NJ: Lawrence Erlbaum.

Blunch, N. J. (2008). *Introduction to Structural Equation Modelling: Using SPSS and AMOS.*

Boomsma, A. (2000). Reporting analyses of covariance structures. *Structural Equation Modelling*, **7**, 461-483.

Branden, N. (1969). The psychology of self-esteem. New York: Bantam

Browne, M. W., & Cudeck, R. (1993). Alternative ways of assessing model fit. In K. A. Bollen & J. S. Long (Eds.), *Testing structural equation models* (pp. 136-162). Newbury Park, CA: Sage.

Bull, S. J., Albinson, J. G., & Shambrook, C. J. (1996). *The mental game plan: Getting psyched for sport.* Eastbourne: Sports Dynamics.

Bull, S. J., Shambrook, C. J., James, W., & Brooks, J. E. (2005). Towards an understanding of mental toughness in elite English cricketers. *Journal of Applied Sport Psychology*, **17**, 209-227.

Burisch, M. (1997). Test length and validity revisited. *European Journal of Personality*, **11**, 303-315.

Byrne, B. M. (2010). *Structural equation modelling with AMOS: basic concepts, applications, and programming* (2nd ed.). New York: Routledge.

Campbell, D. T., & Stanley, J. C. (1966). Experimental and quasi experimental designs for research. Chicago: Rand McNally.

Carron, A. V., Bray, S. R., & Eys, M. A. (2002). Team cohesion and team success in sport. *Journal of Sport Sciences*, **20**, 119-126.

Carver, C. S. (1989). How should multifaceted personality constructs be tested? Issues illustrated by self-monitoring, attributional style, and hardiness. *Journal of Personality and Social Psychology*, **56**, 577-585.

Chelladurai, P., & Saleh, S. (1978). Preferred leadership in sports. *Canadian Journal of Applied Sport Sciences*, **3**, 85-92.

Chen, F. F., Sousa, K. H., & West, S. G. (2005). Testing measurement invariance of second-order factor models. *Structural Equation Modeling*, **12**, 471–492.

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

Churchill, G. A., Jr. (1979). A paradigm for developing better measures of marketing constructs. *Journal of Marketing Research*, **16**, 64-73.

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

Clough, P., Earle, K., & Sewell, D. (2002). Mental toughness: The concept and its measurement. In I. Cockerill (Eds.), *Solutions in Sport Psychology* (pp. 32-45). London: Thomson.

Clough, P., Marchant, D., & Earle, K. (2007). *The MTQ48 Technical Manual.* Retrieved from http://www.aqr.co.uk/sites/default/files/MTQ48%20Technical%20Manual%20Jan%202007.pdf.

Cohen, S., Kamarck, T., & Mermelstein, R. (1983). A global measure of perceived stress. *Journal of Health and Social Behavior,* **24**, 386-396.

Comrey, A. L., & Lee, H. B. (1992). *A first course in Factor Analysis* (2nd ed.). Mahwah, NJ: Lawrence Erlbaum Associates.

Connaughton, D., & Hanton, S. (2009). Mental toughness in sport: Conceptual and practical issues. In S. D. Mellalieu & S. Hanton (Eds.), *Advances in applied sport psychology: A review* (pp. 317-346). London: Routledge.

Connaughton, D., Thelwell, R., & Hanton, S. (2012). Mental toughness development: Issues, practical implications and future directions. In S. Gordon & D. F. Gucciardi (Eds.), *Mental toughness in sport: Developments in theory and research* (pp. 135-162). London: Routledge.

Connor, K., & Davidson, J. (2003). Development of a new resilience scale: The Connor-Davidson Resilience Scale (CD-RISC). *Depression and Anxiety,* **18**, 76–82.

Conroy, D. E., Elliot, A. J., & Hofer, S. M. (2003). A 2x2 Achievement Goals Questionnaire for Sport: Evidence for factorial invariance, temporal stability, and external validity. *Journal of Sport and Exercise Psychology,* **25**, 456-476.

Cooper, C. (2010). *Personality and Individual differences* (3rd ed.). London, UK: Hodder Education.

Coulter, T. J., Mallett, C. J., & Gucciardi, D. F. (2010). Understanding mental toughness in Australian soccer: Perceptions of players, parents, and coaches. *Journal of Sports Sciences,* **28**, 699-716.

Cronbach, L. J., & P. E. Meehl (1955). Construct validity in psychological tests. *Psychological Bulletin,* **52**, 281-302.

Crust, L. (2009). The relationship between mental toughness and affect intensity. *Personality and Individual Differences,* **47**, 959-963.

Crust, L., & Azadi, K. (2009). Leadership preferences of mentally tough athletes. *Personality and Individual Differences,* **47**, 326-330.

Crust, L., & Azadi, K. (2010). Mental toughness and athletes' use of psychological strategies. *European Journal of Sport Science,* **10**, 43-51.

Crust, L., & Clough, P. J. (2005). Relationship between mental toughness and physical endurance. *Perceptual and Motor Skills,* **100**, 192-194.

Crust, L., & Keegan, R. (2010). Mental toughness and attitudes to risk-taking. *Personality and Individual Differences*, **49**, 164-168.

Crust, L., & Swann, C. (2011a). The relationship between mental toughness and dispositional Flow. *European Journal of Sport Science,* **13**, 215-220.

Crust, L., & Swann, C. (2011b). Comparing two measures of mental toughness. *Personality and Individual Differences,* **50**, 217-221.

Csikzentmihalyi, M (1975). *Beyond boredom and anxiety.* San Francisco: Josey-Bass.

Cumming, J., Clark, S. E., Ste-Marie, D. M., McCullagh, P., & Hall, C. (2005). The Functions of Observational Learning Questionnaire. *Psychology of Sport and Exercise,* **6**, 517-537.

Curran, P. J., West, S. G., & Finch, J. F. (1996). The robustness of test statistics to nonnormality and specification error in confirmatory factor analysis. *Psychological Methods*, **1**, 16–29.

Deaner, H., & Silva, J. M. (2002). Personality and sport performance. In J. M. Silva & D. E. Stevens (Eds.), *Psychological Foundations of Sport* (pp. 48-65). Boston, MA: Allyn and Bacon.

DeVellis, R. F. (1991). *Scale Development: Theory and applications,* Newbury Park, CA: Sage Publications.

Dietrich, H., & Ehrlenspiel, F. (2010). Cognitive interviewing: A qualitative tool for improving questionnaires in sport science. *Measurement in Physical Education and Exercise Science,* **14**, 51–60.

Dolan, C. V. (1994). Factor analysis of variables with 2, 3, 5 and 7 response categories: A comparison of categorical variable estimators using simulated data. *British Journal of Mathematical and Statistical Psychology*, **47**, 309-326.

Eagleton, J. R., McKelvie, S. J., & De Man, A. (2007). Extraversion and neuroticism in team sport participants, individual sport participants, and nonparticipants. *Perceptual & Motor Skills*, **105**, 265-275.

Earle, K. (2006). *The construct of mental toughness: A psychometric and experimental analysis.* Unpublished PhD thesis, University of Hull.

Egloff, B., & Gruhn, A. J. (1996). Personality and endurance sports. *Personality and Individual Differences,* **21**, 223-229.

Eid, J., & Morgan, C.A. (2006). Dissociation, hardiness, and performance in military cadets participating in survival training. *Military Medicine*, **171**, 436-42.

Evans, V., & Quarterman, J. (1983). Personality characteristics of successful and unsuccessful black female basketball players. *International Journal of Sports Psychology*, **14**, 105-115.

Fallowfield, J. L., Hale, B. J., & Wilkinson, D. M. (2005). *Using Statistics in Sport and Exercise Science Research.* Chichester: Lotus Publishing.

Favret, B., & Benzel, D. (1997). *Complete guide to water skiing.* Champaign, IL: Human Kinetics.

Florian, V., Mikulincer, M., & Taubman, O. (1995). Does hardiness contribute to mental health during a stressful real-life situation: The roles of appraisal and coping. *Journal of Personality and Social Psychology*, **68**, 687-695.

Folkman, S., & Moskowitz, J. T. (2004). Coping: pitfalls and promise. *Annual Review of Psychology*, **55**, 745-74.

Ford, I. W., Eklund, R. C., & Gordon, S. (2000). An examination of psychosocial variables moderating the relationship between life stress and injury time-loss among athletes of a high standard. *Journal of Sports Sciences*, **18**, 301-312.

Fornell, C., & Larcker, D. F. (1981). Evaluating structural with unobservable variables and measurement error. *Journal of Marketing Research,* **18,** 39-50.

Fourie, S., & Potgieter, J. R. (2001). The nature of mental toughness in sport. *South African Journal for Research in Sport, Physical Education & Recreation,* **23**, 63-72.

Franken, R., Gibson, K., & Rowland, G. (1992). Sensation seeking and the tendency to view the world as threatening. *Personality and Individual Differences,* **13**, 31-38.

Frankl, V. E. (1963). *Man's search for meaning: An introduction to logotherapy* (L. Lasch, Trans.). New York: Washington Square Press. (Original work published 1946).

Franklin, S. B., Gibson, D. J., Robertson, P. A., Pohlmann, J. T., & Fralish, J. S. (1995). Parallel Analysis: a method for determining significant principle components. *Journal of Vegetation Science,* **6**, 99-106.

Franks, D. D., & Marolla, J. (1976). Efficacious action and social approval as interacting dimensions of self-esteem: A tentative formulation through construct validation. *Sociometry*, **39**, 324-341.

Freeman, P., Coffee, P., & Rees, T. (2011). The PASS-Q: The perceived available support in sport questionnaire. *Journal of Sport & Exercise Psychology*, **33**, 54-74.

Funk, S. C., & Houston, B. K. (1987). A critical analysis of the hardiness scale's validity and utility. *Journal of Personality and Social Psychology*, **53**, 572-578.

Garson, G. D. (2006). *Factor analysis. Statnotes: Topics in multivariate analysis.* Retrieved from N. Carolina St. University, Quantitative research in Public Administration Web site: http://www2.chass.ncus.edu/garson/pa765/factor.htm.

Gee, C. J., Marshall, J. C., & King, J. F. (2010). Should coaches use personality assessments in the talent identification process? A 15 year predictive study on professional hockey players. *International Journal of Coaching Science*, **4**, 25-34.

Gendlin, E. T. (1966). Existentialism and experiential psychotherapy. In C. Moustakas (Eds.), *The child's discovery of himself* (pp. 196–236). New York: Basic Books.

Gerbing, D. W., & J. G. Hamilton (1996). Viability of exploratory factor analysis as a precursor to confirmatory factor analysis. *Structural Equation Modeling*, **3**, 62-72.

Gignac, G. E. (2009). Psychometrics and the measurement of emotional intelligence. In C. Stough, D. H. Saklofske & J. D. A. Parker (Eds.), *Assessing emotional intelligence: Theory, research, and applications* (pp. 9-42), New York: Springer.

Golby, J., & Sheard, M. (2004). Mental toughness and hardiness at different levels of rugby league. *Personality and Individual Differences*, **37**, 933-942.

Goldberg, A. S. (1998). *Sports slump busting: 10 steps to mental toughness and peak performance.* Champaign, IL: Human Kinetics.

Goss, J. D. (1994). Hardiness and mood disturbances in swimmers while over training. *Journal of Sport and Exercise Psychology*, **16**, 135-149.

Gould, D., Hodge, K., Peterson, K., & Petlichkoff, L. (1987). Psychological foundations of coaching: Similarities and difference among intercollegiate wrestling coaches. *The Sport Psychologist*, **1**, 293-308.

Green, S. (1991). How many subjects does it take to do a regression analysis? *Multivariate Behavioral Research*, **26**, 499-510.

Greenier, K. D., Kernis, M. H., & Waschull, S. B. (1995). Not all high (or low) self-esteem people are the same: Theory and research on stability of self-esteem. In M. H. Kernis (Eds.), *Efficacy, agency, and self -esteem* (pp. 51-71). New York: Plenum.

Greenwood, M., & Simpson, W. K. (1994). Personality traits of intercollegiate baseball athletes playing central versus noncentral defensive positions at three levels of competition. *Applied Research in Coaching and Athletics International*, **1**, 15-30.

Gucciardi, D. F. (2009). Do developmental differences in mental toughness exist between specialised and invested Australian footballers? *Personality and Individual Differences,* **47**, 985-989.

Gucciardi, D. F., & Gordon, S. (2009). Development and preliminary validation of the Cricket Mental Toughness Inventory (CMTI). *Journal of Sports Sciences*, **27**, 1293-1310.

Gucciardi, D. F., Gordon, S., & Dimmock, J. A. (2008). Towards an understanding of mental toughness in Australian football. *Journal of Applied Sport Psychology,* **20**, 261-281.

Gucciardi, D. F., Hanton, S., Mallett, C. J. (2012). Progressing measurement in mental toughness: A case example of the Mental Toughness Questionnaire 48. *Sport, Exercise, and Performance Psychology,* **1**, 194-214.

Gucciardi, D. F., Mallett, C. J., Hanrahan, S. J., & Gordon, S. (2011). Measuring mental toughness in sport. In S. Gordon & D. F. Gucciardi (Eds.), *Mental Toughness in Sport: Developments in theory and research* (pp. 108-132). London: Routledge.

Gucciardi, D., Gordon, S., & Dimmock, J. (2009). Development and preliminary validation of a mental toughness inventory for Australian football. *Psychology of Sport and Exercise,* **10**, 201-209.

Hagger, M. S., & Chatzisarantis, N. L. D. (2009). Assumptions in research in sport and exercise psychology. *Psychology of Sport & Exercise*, **10**, 511-519.

Hall, C., Mack, D., Paivio, A., & Hausenblas, H. (1998). Imagery use by athletes: Development of the sport imagery questionnaire. *International Journal of Sport Psychology*, **29**, 73-89.

Hanton, S., Evans, L., Neil, R. (2003). Hardiness and the competitive trait anxiety response. *Anxiety, Stress, and Coping*, **16,** 167-184.

Hanton, S., Neil, R., & Mellalieu, S. D. (2008). Recent developments in competitive anxiety direction and competition stress research. *International Review of Sport and Exercise Psychology,* **1**, 45-57.

Hanton, S., Neil, R., Mellalieu, S. D., & Fletcher, D. (2008). Competitive experience and performance status: An investigation into multidimensional anxiety and coping. *European Journal of Sports Science,* **8,** 143-152.

Hardy, L. (1996). A test of the catastrophe models of anxiety and sports performance against Multidimensional Anxiety Theory models using the method of dynamic differences. *Anxiety, Stress, & Coping,* **9**, 69-86.

Hardy, L., Woodman, T., & Carrington, S. (2004). Is self-confidence a bias factor on higher-order catastrophe models? An exploratory analysis. *Journal of Sport and Exercise Psychology,* **26**, 359-368.

Harter, S. (1999). *The construction of the self: A developmental perspective*. New York: Guilford Press.

Hinkin, T. R. (1995). A review of scale development practices in the study of organisations. *Journal of Management,* **21**, 967-988.

Hodge, K., Lonsdale, C., & Jackson, S. A. (2009). Athlete engagement in elite sport: An exploratory investigation of antecedents and consequences. *The Sport Psychologist*, **23**, 186-202.

Hogan, R. (1998). Reinventing personality. *Journal of Social and Clinical Psychology*, **17**, 1-10.

Hogan, R., & Shelton, D. (1998). A socioanalytic perspective on job performance. *Human Performance*, **11**, 129-144.

Horn, J. L., & McArdle, J. J. (1992). A practical and theoretical guide to measurement invariance in aging research. *Experimental Aging Research*, **18**, 117-144.

Horsburgh, V. A., Schermer, J. A., Veselka, L., & Vernon, P. A. (2009). A behavioural genetic study of mental toughness and personality. *Personality and Individual Differences*, **46**, 100-105.

Hoyle, R. H., & Panter, A. T. (1995). Writing about structural equation models. In R. H. Hoyle (Eds.), *Structural equation modelling: Concepts, issues, and applications* (pp. 158-176). Thousand Oaks, CA: Sage.

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

Hu, L., Bentler, P. M., & Kano, Y. (1992). Can test statistics in covariance structure analysis be trusted? *Psychological Bulletin*, **112**, 351-362.

Hull, J. G., Van Treuren, R. R., & Virnelli, S. (1987). Hardiness and health: A critique and alternative approach. *Journal of Personality and Social Psychology*, **53**, 518-530.

Hurley, A. E., Scandura, T. A., Schriesheim, C. A., Brannick, M. T., Seers, A., Vandenberg, R. J., & Williams, L. J. (1997). Exploratory and confirmatory factor analysis: Guidelines, issues, and alternatives. *Journal of Organizational Behavior*, **18**, 667-683.

Hutcheson, G., & Sofroniou, N. (1999). *The Multivariate Social Scientist*. Sage Publications Ltd: London.

Hystad, S. W., Eid, J., Johnsen, B. H., Laberg, J. C., & Bartone, P. T. (2010). Psychometric properties of the revised Norwegian dispositional resilience (hardiness) scale. *Scandinavian Journal of Psychology,* **51**, 237-245.

International Test Commission (1999). *International guidelines for test use.* Retrieved October 31, 2012 from http://www.intestcom.org/upload/sitefiles/41.pdf.

Jackson, D. A. (1993). Stopping rules in principal components analysis: a comparison of heuristical and statistical approaches. *Ecology*, **74**, 2204-2214.

Jackson, S., & Eklund, R. (2002). Assessing flow in physical activity: The Flow State Scale-2 and Dispositional Flow Scale-2. *Journal of Sport and Exercise Psychology,* **24**, 133-150.

Jackson, S., & Kimiecik, J. (2008). The flow perspective of optimal experience in sport and physical activity. In T. Horn (Eds.), *Advances in Sport Psychology* (pp. 377-400). Champaign, IL: Human Kinetics.

Jones, G. (1995). More than just a game: Research developments and issues in competitive anxiety in sport. *British Journal of Psychology,* **86**, 449-478.

Jones, G., & Hanton, S. (2001). Pre-competitive feeling states and directional anxiety interpretations. *Journal of Sports Sciences*, **19**, 385-395.

Jones, G., Hanton, S., & Connaughton, D. (2002). What is this thing called mental toughness? An investigation of elite sport performers. *Journal of Applied Sport Psychology,* **14**, 205-218.

Jones, G., Hanton, S., & Connaughton, D. (2007). A framework of mental toughness in the world's best performers. *The Sport Psychologist,* **21**, 243-264.

Jones, J. G., Hanton. S., & Swain. A. (1994). Intensity and interpretation of anxiety symptoms in elite and non-elite sports performers. *Personality and Individual Differences,* **17**, 657-663.

Jones, J. G., & Swain, A. (1995). Predispositions to experience debilitative and facilitative anxiety in elite and non-elite performers. *The Sport Psychologist,* **9**, 201-211.

Jones, M. V., Meijen, C., McCarthy, P. J., & Sheffield, D. (2009). A theory of challenge and threat states in athletes. *International Review of Sport and Exercise Psychology*, **2**, 161–180.

Jöreskog, K. G., & Sörbom, D. (1993). *LISREL8: User's Reference Guide.* Chicago: Scientific Software International.

Jöreskog, K. G., & Sörbom, D. (1996). *LISREL 8 User's Reference Guide.* Chicago, IL: Scientific Software International.

Judkins, S. K., Arris, L., & Keener, E. (2005). Program evaluation in graduate nursing education: Hardiness as a predictor of success among nursing administration students. *Journal of Professional Nursing.* **21**, 314-321.

Judkins, S., & Rind, R. (2005). Hardiness, job satisfaction, and stress among home health nurses. *Home Health Care Management & Practice,* **17**, 113-118.

Kais, K., & Raudsepp, L. (2004). Cognitive and somatic anxiety and self-confidence in athletic performance of beach volleyball. *Perceptual & Motor Skills,* **98**, 439-449.

Kaiseler, M. K., Polman, R. C. J., & Nicholls, A. R. (2009). Mental toughness, stress, stress appraisal, coping, and coping effectiveness. *Personality and Individual Differences,* **47***,* 728-733.

Kelley, B. C. (1994). A model of stress and burnout in collegiate coaches: Effects of gender and time of season. *Research Quarterly for Exercise and Sport*, **65**, 48-58.

Kelley, B. C., Eklund, R. C., & Ritter-Taylor, M. (1999). Stress and burnout among collegiate tennis coaches. *Journal of Sport and Exercise Psychology*, **21**, 113-130.

Kelloway, K. E. (1995). Structural equation modelling in perspective. *Journal of Organizational Behavior,* **16**, 215-224.

Kelley, T. L. (1927). *Interpretation of Educational Measurements*. Yonkers-on-Hudson, N.Y: World Book Company.

Kelly, G. A. (1991). *The psychology of personal constructs; A theory of personality* (vol. 1). London: Routledge; originally published 1955.

Kierkegaard, S. (1954). *Fear and trembling and the sickness unto death.* Garden City, NY: Doubleday Anchor.

Klag, S., & Bradley, G. (2004). The role of hardiness in stress and illness: An exploration of the effect of negative affectivity and gender. *British Journal of Health Psychology*, **9**, 137–161.

Kline, P. (1986). *A handbook of test construction*. London: Methuen.

Kline, P. (1993). *The handbook of psychological testing*. London: Routledge.

Kline, R. B. (1998). *Principles and practice of structural equation modelling.* New York: Guilford Press.

Kline, P. (2000). *The handbook of psychological testing* (2nd ed.). London: Routledge.

Kline, R. B. (2005). *Principles and practice of structural equation modeling* (2nd ed.). New York: The Guilford Press.

Kobasa, S. C. (1979). Stressful life events, personality, and health: An inquiry into hardiness. *Personality and Social Psychology,* **37**, 1-11.

Kobasa, S. C. (1982). Commitment and coping in stress resistance among lawyers. *Journal of Personality and Social Psychology*, **42**, 707-717.

Kobasa, S. C., Maddi, S. R., & Kahn, S. (1982). Hardiness and health: A prospective study. *Journal of Personality and Social Psychology*, **42**, 884-890.

Kowalski, K. C., & Crocker, P. R. E. (2001). The development and validation of the Coping Function Questionnaire for adolescents in sport. *Journal of Sport & Exercise Psychology*, **23**, 136-155.

Lane, A. M., Harwood, C., Terry, P. C., & Karageorghis, C. I. (2004). Confirmatory factor analysis of the Test of Performance Strategies (TOPS) among adolescent athletes. *Journal of Sports Sciences*, **22**, 803-812.

Lane, A. M., Nevill, A. M., Bowes, N., & Fox, K. R. (2005). Test-retest stability of the task and ego orientation questionnaire. *Research Quarterly for Exercise and Sport*, **76**, 339-346.

Larsen, R. J. (1984). Theory and measurement of affect intensity as an individual difference characteristic. *Dissertation Abstracts International,* **5**, 2297B.

Leary, M. R., & Downs, D. L. (1995). Interpersonal functions of the self-esteem motive: The self-esteem system as a sociometer. In M. H. Kernis (Eds.), *Efficacy, agency, and self-esteem* (pp. 123-144). New York: Plenum.

Llewellyn, D., & Sanchez, X. (2008). Individual differences and risk taking in rock climbing. *Psychology of Sport and Exercise,* **9**, 413-426.

Loehr, J. E. (1982). *Athletic excellence: mental toughness training for sports.* New York: Plume.

Loehr, J. E. (1986). *Mental toughness training for sports; achieving athletic excellence.* Lexington, MA: Stephen Greene Press.

Loehr, J. E. (1995). *The new toughness training for sports.* New York: Penguin.

Lonsdale, C., Hodge, K., & Rose, E. A. (2006). Pixels vs. paper: Comparing online and traditional survey methods in sport psychology. *Journal of Sport and Exercise Psychology*, **28**, 100-108.

Lonsdale, C., Hodge, K., & Rose, E. A. (2008). The development of the Behavioral Regulation in Sport Questionnaire (BRSQ): Instrument development and initial validity evidence. *Journal of Sport and Exercise Psychology*, **30**, 323-335.

Luszki, W. A. (1982). *Winning tennis through mental toughness.* New York: Everest House.

Lynn, M. R. (1986). Determination and quantification of content validity. *Nursing Research,* **35**, 382-385.

MacKenzie, S. B. (2003). The dangers of poor construct conceptualisation. *Journal of the Academy of Marketing Science,* **31**, 323-326.

MacKenzie, S. B., Podsakoff, P. M., & Podsakoff, N, P. (2011). Construct measurement and validation procedures in MIS and behavioral research: Integrating new and existing techniques. *MIS Quarterly,* **35**, 293-334.

Maddi, S. R. (1990). Issues and interventions in stress mastery. In H. S. Friedman (Eds.), *Personality and disease*. New York: Merrill.

Maddi, S.R. (1997). Personal Views Survey II. In C. P Zalaquett & R. J. Wood (Eds.), *Evaluating stress: A book of resources* (pp. 293-309). Lanham, MD: Scarecrow Press, Inc.

Maddi, S. R. (2002). The story of hardiness: Twenty years of theorizing, research, and practice. *Consulting Psychology Journal,* **54**, 175-185.

Maddi, S. R. (2004). Hardiness: An operationalisation of existential courage. *Journal of Humanistic Psychology*, **44**, 279-298.

Maddi, S. R. (2008). The courage and strategies of hardiness as helpful in growing despite major, disruptive stresses. *American Psychologist*, **63**, 563-564.

Maddi, S. R., Harvey, R. H., Khoshaba, D. M., Lu, J. L., Persico, M., & Brow, M. (2006). The personality construct of hardiness, III: Relationships with repression, innovativeness, authoritarianism, and performance. *Journal of Personality*, **74**, 575-597.

Maddi., S. R., & Hess, M. J. (1992). Personality hardiness and success in basketball. *International Journal of Sport Psychology*, **23**, 360-368.

Maddi, S. R., & Hightower, M. (1999). Hardiness and optimism as expressed in coping patterns. *Consulting Psychology Journal*, **51**, 95-105.

Maddi, S. R., Hoover, M., & Kobasa, S. C. (1982). Alienation and exploratory behaviour. *Journal of Personality and Social Psychology*, **42**, 884-890.

Maddi, S. R., & Khoshaba, D. M. (2001). Personal views survey (3rd ed.). Newport Beach, CA: The Hardiness Institute.

Maddi, S. R., & Kobasa, S. C. (1984). *The hardy executive: Health under stress.* Homewood, IL: Dow-Jones Irwin.

Maddi, S. R., Propst, B. S., & Feldinger, I. (1965). Three expressions of the need for variety. *Journal of Personality*, **19,** 73-76.

Marchant, D. C., Polman, R. C. J., Clough, P. J., Jackson, J. G., Levy, A. R., & Nicholls, A. R. (2009). Mental toughness in the workplace: Managerial and age differences. *Journal of Managerial Psychology,* **24**, 428-437.

Mardia, K. V. (1970). Measures of multivariate skewness and kurtosis with applications. *Biometrika*, **57**, 519–530.

Mardia, K. V. (1974). Applications of some measures of multivariate skewness and kurtosis for testing normality and robustness studies, *Sankhya*, Series B, **36**, 115-128.

Marsh, H. W. (1997). The measurement of physical self-concept: A construct validation approach. In K. Fox (Eds.), *The physical self: From motivation to well-being* (pp. 27-58). Champaign, IL: Human Kinetics.

Marsh, H. W. (2002). A multidimensional physical self-concept: A construct validity approach to theory, measurement, and research. *Psychology: The Journal of the Hellenic Psychology Society,* **9**, 459-93.

Marsh, H. W. (2007). Application of confirmatory factor analysis and structural equation modeling in sport/exercise psychology. In G. Tenenbaum & R. C. Eklund (Eds.), *Handbook of sport psychology* (3rd ed., pp. 774-798). New York: Wiley.

Marsh, H. W., & Hau, K-T. (2007). Applications of latent-variable models in educational psychology: The need for methodological-substantive synergies. *Contemporary Educational Psychology*, **32**, 151-170.

Marsh, H. W., Hau, K-T., & Wen, Z. (2004). In search of golden rules: comment on hypothesis-testing approaches to setting cutoff values for fit indexes and dangers in overgeneralizing Hu and Bentler's (1999) findings. *Structural Equation Modeling*, **11**, 320-341.

Marsh, H. W., Martin, A. J., & Jackson, S. A. (2010). Introducing a short version of the Physical Self-Description Questionnaire: New strategies, short form evaluation evaluative criteria, and applications of factor analyses. *Journal of Sport and Exercise Psychology*, **32**, 438-482.

Martens, M. P., & Webber, S. N. (2002). Psychometric properties of the sport motivation scale: An evaluation with college varsity athletes from the US. *Journal of Sport and Exercise Psychology*, **24**, 254-270.

Martens, R., Burton, D., Vealey, R. S., Bump, L. A., & Smith, D. E. (1990). Development and validation of the Competitive State Anxiety Inventory-2. In R. Martens, R. S. Vealey & D. Burton (Eds.), *Competitive anxiety in sport* (pp. 117-190). Champaign, IL: Human Kinetics.

Martin, J. J., Kelley, B. C., & Dias, C. (1999). Stress and burnout in female high school athletic directors. *Women in Sport and Physical Activity Journal*, **8**, 101-116.

Martin, J. J., Kelley, B. C., & Eklund, R. C. (1999). A model of stress and burnout in male high school athletic directors. *Journal of Sport and Exercise Psychology*, **21**, 280-294.

Matthews, G., & Deary, I. J. (1998). *Personality traits*. Cambridge: Cambridge University Press.

McCrae, R. R., & Costa, P. T. (2004). A contemplated revision of the NEO Five-Factor Inventory. *Personality and Individual Differences*, **36**, 587-596.

McCrae, R. R., Costa, P. T., Jr., Ostendorf, F., Angleitner, A., Hrebícková, M., Avia, M. D., Sanz, J., Sanchez-Bernardos, M. L., Kusdil, M. E., Woodfield, R. Saunders, P. R., Smith, P. B. (2000). Nature over nurture: Temperament, personality, and life span development. *Journal of Personality and Social Psychology*, **78**, 173-186.

McCrae, R. R., Kurtz, J. E., Yamagata, S., & Terracciano, A. (2011). Internal consistency, retest reliability, and their implications for personality scale validity. *Personality and Social Psychology Review*, **15**, 28–50.

McGrath, R. E. (2005). Conceptual complexity and construct validity. *Journal of Personality Assessment*, **85**, 112–24.

McGregor, H. A., & Elliot, A. J. (2002). Achievement goals as predictors of achievement relevant processes prior to task engagement. *Journal of Educational Psychology*, **94**, 381-395.

McKelvie, S. J., Lemieux, P., & Stout, D. (2003). Extraversion and neuroticism in contact athletes, non-contact athletes, and non-athletes: A research note. *Athletic Insight*, **5**, 19-27.

Meade, A. W. (2005, April). *Sample Size and Tests of Measurement Invariance*. Paper presented at the 20th Annual Conference of the Society for Industrial and Organizational Psychology, Los Angeles, CA.

Menard, S. (1995). Applied logistic regression analysis. Sage University paper series on quantitative applications in the social sciences, pp. 7-106. Thousand Oaks, CA. Sage.

Middleton, S. C. (2005). *Mental Toughness: Conceptualisation and measurement.* Unpublished Doctoral Thesis, University of Western Sydney.

Moore, L. J., Vine, S. J., Wilson, M. R., & Freeman, P. (2012). The effect of challenge and threat states on performance: An examination of potential mechanisms. *Psychophysiology*, **49**, 1417-1425.

Morgan, W. P., & Johnson, R. W. (1978). Personality characteristics of successful and unsuccessful oarsmen. *International Journal of Sport Psychology*, **9**, 119-133.

Morgan, W. P., O’Connor, P. J., Ellickson, K. A., & Bradley, P. W. (1988). Personality structure, mood states, and performance in elite distance runners. *International Journal of Sport Psychology*, **19**, 247-263.

Morris, T. (1995). Psychological characteristics and sport behaviour. In T. Morris & J. Summers (Eds.), *Sport psychology: Theory, applications and issues* (pp. 3-27). Brisbane: John Wiley & Sons.

Mruk, C. J. (1995). *Self-esteem: Research, theory, and practice*. New York, NY: Springer Publishing Co, Inc.

Muthén, B., & Kaplan, D. (1985). A comparison of some methodologies for the factor analysis of non-normal Likert variables. *British Journal of Mathematical and Statistical Psychology*, **38**, 171-189.

Myers, R. (1990). Classical and modern regression with applications (2nd ed.). Boston, MA. Duxbury.

Nevill, A. M. (1996). Validity and measurement agreement in sports performance (Editorial). *Journal of Sports Sciences,* **14**, 199.

Nevill, A. M., Lane, A. M., Kilgour, L. J., Bowes, N., & Whyte, G. P. (2001). Stability of psychometric questionnaires. *Journal of Sports Sciences,* **19,** 273-278.

Newcombe, P.A., & Boyle, G. J. (1995). High school students’ sport personalities: Variations across participation level, gender, type of sport, and success. *International Journal of Sport Psychology*, **19**, 247-263.

Nicholls, A. R., & Polman, R. C. J. (2007). Coping in sport: A systematic review. *Journal of Sports Sciences*, **25**, 11-31.

Nicholls, A. R., Levy, A. R., Polman, R. C. J., & Crust, L. (2011). Mental toughness, coping self-efficacy, and coping effectiveness among athletes. *International Journal of Sport Psychology*, **42**, 513-524.

Nicholls, A. R., Polman, R. C. J., Levy, A. R., & Backhouse, S. H. (2008). Mental toughness, optimism, pessimism, and coping among athletes. *Personality and Individual Differences*, **44**, 1182-1192.

Nicholls, A. R., Polman, R. C. J., Levy, A. R., & Backhouse, S. H. (2009). Mental toughness in sport: Achievement level, gender, age, experience, and sport type differences. *Personality and Individual Differences,* **47**, 73-75.

Ntoumanis, N. (2001). A step-by-step guide to SPSS for sport and exercise studies. London: Routledge.

Ntoumanis, N., & Vazou, S. (2005). Peer motivational climate in youth sport: Measurement development and validation. *Journal of Sport & Exercise Psychology*, **27**, 432-455.

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

O'Brien, E., & Epstein, S. (1988). *The Multidimensional Self-Esteem Inventory: Professional manual*. Odessa, FL: Psychological Assessment Resources.

Olsson, U. H., Foss, T., Troye, S. V., & Howell, R. D. (2000). The performance of ML, GLS, and WLS estimation in structural equation modeling under conditions of misspecification and nonnormality. *Structural Equation Modeling,* **7**, 557-595.

Parfitt, G., & Pates, J. (1999). The effects of cognitive and somatic anxiety and self-confidence on components of performance during competition. *Journal of Sports Sciences*, **17**, 351-356.

Peacock, E. J., & Wong, P. T. P. (1990). The Stress Appraisal Measure (SAM): A multidimensional approach to cognitive appraisal. *Stress Medicine*, **6**, 227-236.

Pervin, L., A. (1996). *The science of personality*. New York: John Wiley & Sons.

Peterson, R. A. (2000). Constructing effective questionnaires. Thousand Oaks, CA: Sage.

Phares, E. J. (1976). *Locus of control in personality*. Momstown: General Learning Press.

Piedmont, R. L., Hill, D. C., & Blanco, S. (1999). Predicting athletic performance using the five factor model of personality. *Personality and Individual Differences*, **27**, 769-777.

Podsakoff, P. M., MacKenzie, S. B., Lee, J. Y., & Podsakoff, N. P. (2003). Common method biases in behavioural research: A critical review of the literature and recommended remedies. *Journal of Applied Psychology*, **88**, 879-903.

Prapavessis, H., & Grove, J. R. (1994). Personality variables as antecedents of precompetitive mood state temporal patterning. *International Journal of Sport Psychology*, **22**, 347-365.

Ptacek, J. T., Smith, R. E., Espe, K., & Raffety, B. (1994). Limited correspondence between daily coping reports and retrospective coping recall. *Psychological Assessment*, **6**, 41-49.

Raykov, T., & Marcoulides, G. A. (2000). *A first course in structural equation modeling*, Mahwah, NJ: Lawrence Erlbaum Associates.

Reiss, S., Wiltz, J., & Sherman, M. (2001). Trait motivational correlates of athleticism. *Personality and Individual Differences*, **30**, 1139-1145.

Reynolds, W. M. (1982). Development of reliable and valid short forms of the Marlowe-Crowne Social Desirability Scale. *Journal of Clinical Psychology,* **38**, 119-125.

Rosenberg, M. (1989). *Society and the adolescent self-image*. (Revised ed.). Middeltown, CT: Wesleyan University Press.

Rotter, J. B. (1966). Generalized expectancies for internal versus external control of reinforcement. *Psychological Monographs*, **80** (Whole No. 609).

Russell, D. W. (2002). In search of underlying dimensions: the use (and abuse) of factor analysis in personality and social psychology bulletin. *Personality and Social Psychology Bulletin*, **28**, 1629-1646.

Sartori, G. (1984). Guidelines for concept analysis. In G. Sartori (Eds.), *In social science concepts: A systematic analysis* (pp. 15-85). Beverly Hills, CA: Sage Publications.

Schutz, R. W., & Gessaaroli, M. E. (1993). Use, misuse, and disuse of psychometrics in sport psychology research. In R. N. Singer, M. Murphy & L. K. Tennant (Eds.), *Handbook of research on sport psychology* (pp. 901-917). New York: Macmillan.

Schwarzer, R., & Jerusalem, M. (1995). Generalized Self-Efficacy scale. In J. Weinman, S. Wright & M. Johnston (Eds.), *Measures in health psychology: A user's portfolio. Causal and control beliefs* (pp. 35-37). Windsor, England: NFER-NELSON.

Seligman, M. E. P. (1975). *Helplessness: On depression, development, and death*. San Francisco: W. H. Freeman.

Sheard, M., & Golby, J. (2006). Effect of a psychological skills training programme in swimming performance and positive psychological development. *International Journal of Sport and Exercise Psychology*, **4**, 149-169.

Sheard, M., Golby, J., & van Wersch, A. (2009). Progress toward construct validation of the Sports Mental Toughness Questionnaire. *European Journal of Psychological Assessment*, **25**, 186-193.

Sheard, S. (2009). Hardiness commitment, gender, and age differentiate university academic performance. *British Journal of Educational Psychology*, **79**, 189-204.

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

Simpson, R. J., Gray, S. C., & Florida-James, G. D. (2006). Physiological variables and performance markers of serving soldiers from two "elite" units of the British Army. *Journal of Sports Sciences*, **24**, 597-604.

Singer, R. N. (1969). Personality differences within and between baseball and tennis players. *Research Quarterly*, **40**, 582-588.

Skirka, N. (2000). The relationship of hardiness, sense of coherence, sports participation, and gender to perceived stress and psychological symptoms among college students. *Journal of Sports Medicine and Physical Fitness*, **40**, 63-70.

Smith, N. C., Bellamy, M., Collins, D. J., & Newell, D. (2001). A test of processing efficiency theory in a team sport context. *Journal of Sports Sciences*, **19**, 321-332.

Smith, R. E., Leffingwell, T. R., & Ptacek, J. T. (1999). Can people remember how they coped? Factors associated with discordance between same-day and retrospective reports. *Journal of Personality and Social Psychology*, **76**, 1050-1061.

Smith, R.E., Smoll, F.L., & Schutz, R.W. (1990). Measurement and correlates of sport-specific cognitive and somatic trait anxiety: The Sport Anxiety Scale. *Anxiety Research*, **2**, 263-280.

Smither, J. W., London, M., & Reilly, R. R. (2005). Does performance improve following multisource feedback? A theoretical model, meta-analysis, and review of empirical findings. *Personnel Psychology*, **58**, 33-66.

Spector, P. (1992). *Summated rating scale construction: An introduction*. Newbury Park, CA: Sage.

Stevens, J. (1996). *Applied multivariate statistics for the social sciences.* Mahwah, NJ: Lawrence Erlbaum.

Stone, A. A., Schwartz, J. E., Neale, J. M., Shiffman, S., Marco, C. A., Hickox, M., & Paty, L. S. (1998). A comparison of coping assessed by ecological momentary assessment and retrospective recall. *Journal of Personality and Social Psychology*, **74**, 1670-1680.

Tabachnick, B. G, & Fidell, L. S. (2001). *Using Multivariate Statistics* (4th ed.). MA: Allyn and Bacon.

Tafarodi, R. W., & Swann, W. B., Jr. (1995). Self-liking and self-competence as dimensions of global self-esteem: Initial validation of a measure. *Journal of Personality Assessment*, **65**, 322-342.

Tafarodi, R. W., & Swann, W. B., Jr. (2001). Two-dimensional self-esteem: Theory and measurement. *Personality and Individual Differences*, **31**, 653-673.

Tanaka, J. S. (1993). Multifaceted conceptions of fit in structural equation models. In K. A. Bollen & J. S. Long (Eds.), *Testing structural equation models* (pp. 10-39). Newbury Park, CA: Sage.

Tapp, J. (1991). Mental toughness. *Referee,* **16**, 44-48.

Thelwell, R., & Maynard, I.W. (1998). Anxiety–performance relationships in cricketers: testing the zone of optimal functioning hypothesis. *Perceptual and Motor Skills*, **87**, 675-689.

Thelwell, R., Weston, N., & Greenlees, I. (2005). Defining and understanding mental toughness in soccer. *Journal of Applied Sport Psychology,* **17**, 326-332.

Thomas, O., Hanton, S., & Jones, G. (2002). An alternative approach to short-form self-report assessment of competitive anxiety: A research note. *International Journal of Sport Psychology*, **33**, 325-336.

Thomas, P. R., Murphy, S. M., & Hardy, L. (1999). Test of performance strategies: Development and preliminary validation of a comprehensive measure of athletes' psychological skills. *Journal of Sports Sciences,* **17**, 697-711.

Thompson, B. (2004). *Exploratory and confirmatory factor analysis: Understanding concepts and applications.* Washington, DC, US: American Psychological Association.

Thorndike, E.L. (1904). An introduction to the theory of mental and social measurements. New York: Teacher's College.

Tillich, P. (1952). *The courage to be.* New Haven: Yale University Press.

Todd, M., Tennen, H., Carney, M. A., Armeli, S., & Affleck, G. (2004). Do we know how we cope? Relating daily coping reports to global and time-limited retrospective assessments. *Journal of Personality and Social Psychology*, **86**, 310-319.

Tourangeau, R., Rips, L. J., & Rasinski, K. (2000). *The Psychology of Survey Response.* Cambridge, UK: Cambridge University Press.

Ullman, J. B. (2001). Structural equation modeling. In B. G. Tabachnick & L. S. Fidell (Eds.), *Using multivariate statistics* (4th ed., pp. 653-771). New York: Allyn Bacon.

Vaughn, B. K., & Daniel, S. R. (2012). Conceptualising validity. In G. Tenenbaum, R. C. Eklund & A. Kamata (Eds.), *Measurement in sport and exercise psychology*, (pp. 33-40). Champaign, IL: Human Kinetics.

Vaughn, B. K., Lee, H., & Kamata, A. (2012). Measurement basics, methods, and issues. In G. Tenenbaum., R. C. Eklund & A. Kamata (Eds.), *Measurement in sport and exercise psychology*, (pp. 25-32). Champaign, IL: Human Kinetics.

Vealey, R. S. (1989). Sport personology: A paradigmatic and methodological analysis. *Journal of Sport and Exercise Psychology*, **11**, 216-235.

Vealey, R. S. (2001). Understanding and enhancing self-confidence in athletes. In R. N. Singer, H. A., Hausenblas & C. M. Janelle (Eds.), *Handbook of sport psychology* (pp. 550-565). New York: John Wiley & Sons.

Wadey, R., Evans, L., Hanton, S., & Neil, R. (2012). An examination of hardiness throughout the sport injury process. *British Journal of Health Psychology,* **17**, 103-128.

Wagnild, G., & Young, H. M. (1993). Development and psychometric evaluation of the resilience scale. *Journal of Nursing Measurement,* **1,** 165-178.

Watson, D. (2004). Stability versus change, dependability versus error: Issues in the assessment of personality over time. *Journal of Research in Personality*, **38,** 319-350.

West, S. G., Fitch, F., & Curran, P. J. (1995). Structural equation models with non-normal variables: Problems and remedies. In Hoyle R. H (Eds.), *Structural equation modelling: Concepts, issues, and applications* (pp. 56-75). Thousand Oaks, CA: Sage.

Wheaton, B., Muthen, B., Alwin, D. E., & Summers, G. E. (1977). Assessing reliability and stability in panel models. In D. R. Heise (Eds.), *Sociological methodology* (pp. 85-136). San Francisco: Jossey-Bass.

Wiebe, D. J. (1991). Hardiness and stress moderation: A test of proposed mechanisms. *Journal of Personality and Social Psychology*, **60**, 89-99.

Williams, S. E., & Cumming, J. (2011). Measuring athlete imagery ability: The Sport Imagery Ability Questionnaire. *Journal of Sport & Exercise Psychology,* **33**, 416-440.

Williams, P. G., Wiebe, D. J., & Smith, T. W. (1992). Coping processes as mediators of the relationship between hardiness and health. *Journal of Behavioral Medicine*, **15**, 237-255.

Wilson, K., & Batterman, A. (1999). Stability of questionnaire items in sport and exercise psychology: Bootstrap limits of agreements. *Journal of Sports Sciences,* **17**, 725-734.

Woodcock, C., Duda, J. L., Cumming, J., Sharp, L., & Holland, M. J. G. (2012). Assessing mental skill and technique use in applied interventions: Recognizing and minimising threats to the psychometric properties of the TOPS. *The Sports Psychologist,* **26**, 1-15.

Wygant, S., & Lindorf, R. (1999): Surveying collegiate Net surfers – Web methodology or mythology. *Quirk's Marketing Research Review*, July.

Zourbanos, N., Hatzigeorgiadis, A., Chroni, S., Theodorakis, Y., & Papaioannou, A. (2009). Automatic Self-Talk Questionnaire for Sports (ASTQS): Development and preliminary validation of a measure identifying the structure of athletes' self-talk. *The Sport Psychologist*, **23**, 233-251.

Zwick, W. R., & Velicer, W. F. (1986). Factors influencing five rules for determining the number of components to retain. *Psychological Bulletin*, **99**, 432-442.

APPENDICES

Study 1 Assessment Instruments and SPSS/AMOS Output

Appendix 1.1
The Mental Toughness Questionnaire-48

Please indicate your response to the following items by **circling one** of the numbers, which have the following meaning;
1 = strongly disagree; **2** = disagree; **3** = neither agree nor disagree; **4** = agree; **5** = strongly agree

Please answer these items carefully, thinking about how you are **generally**. Do not spend too much time on any one item. **ANSWER THE QUESTIONS HONESTLY.**

		◂◂ Disagree				**Agree ▸▸**
1	I usually find something to motivate me	1	2	3	4	5
2	I generally feel in control	1	2	3	4	5
3	I generally feel that I am a worthwhile person	1	2	3	4	5
4	Challenges usually bring out the best in me	1	2	3	4	5
5	When working with other people I am usually quite influential	1	2	3	4	5
6	Unexpected changes to my schedule generally throw me	1	2	3	4	5
7	I don’t usually give up under pressure	1	2	3	4	5
8	I am generally confident in my own abilities	1	2	3	4	5
9	I usually find myself just going through the motions	1	2	3	4	5
10	At times I expect things to go wrong	1	2	3	4	5
11	“I just don’t know where to begin” is a feeling I usually have when presented with several things to do at once	1	2	3	4	5
12	I generally feel that I am in control of what happens in my life	1	2	3	4	5
13	However bad things are, I usually feel they will work out positively in the end	1	2	3	4	5
14	I often wish my life was more predictable	1	2	3	4	5
15	Whenever I try to plan something, unforeseen factors usually seem to wreck it	1	2	3	4	5

1 = strongly disagree; **2** = disagree; **3** = neither agree nor disagree; **4** = agree; **5** = strongly agree

		◂◂ Disagree			**Agree ▸▸**	
16	I generally look on the bright side of life	1	2	3	4	5
17	I usually speak my mind when I have something to say	1	2	3	4	5
18	At times I feel completely useless	1	2	3	4	5
19	I can generally be relied upon to complete the tasks I am given	1	2	3	4	5
20	I usually take charge of a situation when I feel it is appropriate	1	2	3	4	5
21	I generally find it hard to relax	1	2	3	4	5
22	I am easily distracted from tasks that I am involved with	1	2	3	4	5
23	I generally cope well with any problems that occur	1	2	3	4	5
24	I do not usually criticise myself even when things go wrong	1	2	3	4	5
25	I generally try to give 100%	1	2	3	4	5
26	When I am upset or annoyed I usually let others know	1	2	3	4	5
27	I tend to worry about things well before they actually happen	1	2	3	4	5
28	I often feel intimidated in social gatherings	1	2	3	4	5
29	When faced with difficulties I usually give up	1	2	3	4	5
30	I am generally able to react quickly when something unexpected happens	1	2	3	4	5
31	Even when under considerable pressure I usually remain calm	1	2	3	4	5
32	If something can go wrong, it usually will	1	2	3	4	5
33	Things just usually happen to me	1	2	3	4	5
34	I generally hide my emotion from others	1	2	3	4	5
35	I usually find it difficult to make a mental effort when I am tired	1	2	3	4	5

1 = strongly disagree; **2** = disagree; **3** = neither agree nor disagree; **4** = agree; **5** = strongly agree

		◀◀ Disagree				**Agree ▶▶**
36	When I make mistakes I usually let it worry me for days after	1	2	3	4	5
37	When I am feeling tired I find it difficult to get going	1	2	3	4	5
38	I am comfortable telling people what to do	1	2	3	4	5
39	I can normally sustain high levels of mental effort for long periods	1	2	3	4	5
40	I usually look forward to changes in my routine	1	2	3	4	5
41	I feel that what I do tends to make no difference	1	2	3	4	5
42	I usually find it hard to summon enthusiasm for the tasks I have to do	1	2	3	4	5
43	If I feel somebody is wrong, I am not afraid to argue with them	1	2	3	4	5
44	I usually enjoy a challenge	1	2	3	4	5
45	I can usually control my nervousness	1	2	3	4	5
46	In discussions, I tend to back-down even when I feel strongly about something	1	2	3	4	5
47	When I face setbacks I am often unable to persist with my goal	1	2	3	4	5
48	I can usually adapt myself to challenges that come my way	1	2	3	4	5

Permission to use Mental Toughness Questionnaire-48

From: Peter J Clough [P.J.Clough@hull.ac.uk]
Sent: 24 March 2010 15:11
To: Iain Greenlees
Subject: RE: MT-48

All sounds good Iain. Happy for you and your group to use the instrument. Just keep me informed.

Yes, I hold copyright

It would be good to meet up sometime

Peter

From: Iain Greenlees [mailto:I.Greenlees@chi.ac.uk]
Sent: Mon 22/03/2010 15:48
To: Peter J Clough
Subject: MT-48

Dear Peter,

Hello, my name is Iain Greenlees and I am the supervisor of Phil Birch's PhD. I believe that he has been in contact with you concerning using the MT-48. Thank you for the advice and help that you have given him so far and for your willingness to allow us to use the MT-48 in this research. I am just emailing to follow up on a couple of aspects of the permission so that we can proceed in the knowledge that we have covered all bases.

First, I can confirm that Phil and myself (and, if permissible, any other student who expresses an interest in mental toughness research. If this is problematic then we would be more than happy if it was just Phil and myself who had the permission) would only be interested in using the questionnaire for research purposes. The plan for the phd would be to conduct 4 or 5 studies, the first 2 would be reliability/validity studies and the later studies would be aimed at examining predictive validity/behavioural and cognitive differences between high and low mentally tough athletes). If you need this in letter format then please let me know and I will be happy to provide it.

Second, I just wanted to check if you hold the copyright or if we also need to gain the permission of AQR for the purposes of our research. As I said earlier, I do want to ensure that we are covering all bases before the phd begins.

Thanks again for your help with this, I am excited about the possibilities for Phil's phd and am keen to use the MT-48 if possible. It would be good to take you up on your offer of coming down to present on the topic here at some point. I need to find out our guest speaker budget but I hope to be able to invite you down at some point this year if you are still interested.

Best wishes
Iain
Dr Iain Greenlees
Reader in Sport Psychology
Faculty of Sport, Education & Social Sciences
University of Chichester
College Lane
Chichester
West Sussex
PO19 6PE
tel: 01243 816437
email: i.greenlees@chi.ac.uk

Appendix 1.2
The Mental Toughness Questionnaire-48 scoring key

Clough Hardiness Indicator (or Mental Toughness) questionnaire

Sum items and account for reverse scoring of specified items (**R**)

Constructs	**Items**
Challenge	4, 6**R**, 14**R**, 23, 30, 40, 44, 48
Commitment	1, 7, 11**R**, 19, 22**R**, 25, 29**R**, 35**R**, 39, 42**R**, 47**R**
Control- emotion	21**R**, 26**R**, 27**R**, 31, 34, 37**R**, 45
Control-life	2, 5, 9**R**, 12, 15**R**, 33**R**, 41**R**
Confidence-abilities	3, 8, 10**R**, 13, 16, 18**R**, 24, 32**R**, 36**R**
Confidence-interpersonal	17, 20, 28**R**, 38, 43, 46**R**

Total MT score Add all 48 items (accounting for reverse scores!) and divide by 48 to give a mean total MT score

Appendix 1.3
Social Desirability Scale

Please read each statement and circle either true (T) or false (F) that best describes how you are **generally.**

		TRUE	FALSE
1	It is sometimes hard for me to go on with my work if I am not encouraged	T	F
2	I sometimes feel resentful when I don't get my way	T	F
3	There have been times when I feel like rebelling against people in authority even though I knew they were right	T	F
4	No matter who I'm talking to, I'm always a good listener	T	F
5	There have been occasions when I took advantage of someone	T	F
6	I'm always willing to admit it when I make a mistake	T	F
7	I sometimes try to get even rather than forgive and forget	T	F
8	I am always courteous, even to people who are disagreeable	T	F
9	I have never been irked when people expressed ideas very different from my own	T	F
10	There have been times when I was quite jealous of the good fortune of others	T	F
11	I am sometimes irritated by people who ask favours of me	T	F
12	I have never deliberately said something that hurt someone's feelings	T	F

Appendix 1.4
Social Desirability Scale Scoring Key

Negatively worded items = 1, 2, 3, 5, 7, 10, 11.

For positively worded items: True response = 1

False response = 0

For negatively worded items: True response = 0

False response = 1

Sum scores to calculate total Social Desirability score

Appendix 1.5
Ethical approval

Research and Employer Engagement Office
The University of Chichester
College Lane
Chichester
West Sussex PO19 6PE

28 May 2013

To whom it may concern,

Re. Ethical Approval – Phil Birch

Please find below confirmation of the granting of Ethical approval for the following Ethical Review Application submitted by Research Postgraduate student Phil Birch.

Development of the Mental Toughness Questionnaire-48 (8 June 2010).

Yours sincerely

A. Walsh

Antony Walsh
Postgraduate Research Coordinator

Appendix 1.6
Participant information sheet, consent form, and demographic questionnaire

Information Sheet for Participants

PLEASE READ THE FOLLOWING CAREFULLY

Study title: An examination of the psychological characteristics and thought processes of competitive athletes.

What is the study about?
The purpose of this study is to examine the psychological characteristics and thought processes of competitive athletes.

At the University of Chichester before any project starts it has to be checked by university staff. They make sure that the research is good enough to carry out. This project has been checked by the staff of the Department of Sport and Exercise Science and has been passed by the University Ethics Committee.

Do you have to take part?
No! It is your choice whether you participate in the project or not. If at any time during the project you feel that you don't want to continue then you can tell the researcher that you want to stop – you do not have to give a reason.

What will you do in the project?
The study involves completing two questionnaires which will assess your psychological attributes. Please take your time to complete the questionnaires and answer as truthfully as possible.

In total, each questionnaire should take no longer than 6 minutes so the total amount of time you will need to commit to this project is 12 minutes.

Why have you been asked to take part?

You have been asked to take part in this project since we need athletes who are competing in sport to gain an insight into the psychology of sport.

What happens to the information in the project?
All the information which is collected about you during the project will be kept private – no one will know that the information belongs to you and all the information will be kept at the University of Chichester.

Thank you for your time - please ask any questions if you are unsure or confused about what we have said.

What happens if you agree to take part?
- If you are happy to be involved in the project then please sign the informed consent form and the accompanied questionnaires.
- If you do not want to be involved in the project then thank you.

Who can you contact if you have any questions about the project?
Phil Birch – Email p.birch@chi.ac.uk: Phone 01243 816343 or
Dr Iain Greenlees - Email i.greenlees@chi.ac.uk: Phone 01243 816437

Thank you for your time

Consent Form

I understand that my participation in this project will involve completing 2 questionnaires which will assess my perceptions of psychological characteristics and thought processes. I understand that participation in this study is entirely voluntary and that I can withdraw from the study at any time without giving a reason.

I understand that I am free to ask any questions at any time. I am free to withdraw or discuss my concerns with Phil Birch (p.birch@chi.ac.uk) or Dr. Iain Greenlees (i.greenlees@chi.ac.uk).

I understand that the information provided by me will be held totally anonymously, so that it is impossible to trace this information back to me individually. I understand that this information may be retained indefinitely.

I also understand that at the end of the study I will be provided with additional information and feedback about the purpose of the study.

I, ____(NAME) consent to participate in the study conducted by Phil Birch

Signed:

Date:

F Version 1: 26/04/2010

Demographic Questionnaire

Please fill in your details in the spaces provided

Age ______________________________

(years)

Gender Male ☐ Female ☐

Ethnic origin – Please indicate the most appropriate description of your ethnic origin.

(Please tick)

White - British	☐	Asian - Indian	☐
White - Irish	☐	Asian - Pakistani	☐
White – Any other White background	☐	Asian – Bangladeshi	☐
Mixed – White & Black Caribbean	☐	Black or Black British - Caribbean	☐
Mixed – White & Black African	☐	Black or Black British - African	☐
Mixed – White & Black Asian	☐	Black or Black British - Any other Black background	☐
Mixed – Any other Mixed background	☐	Chinese – Any Chinese background	☐
Any other ethnic background	☐		

University of study ____

Year of study ____

(e.g., 1, 2, 3)

What is your primary sport? ____

How many years competitive playing experience do you have in your primary sport? ___________________________

What is the highest level that you have played in your primary sport? ____

F A
Version 1: 26/04/2010

Appendix 1.7
Participant written debrief

Participant Debrief

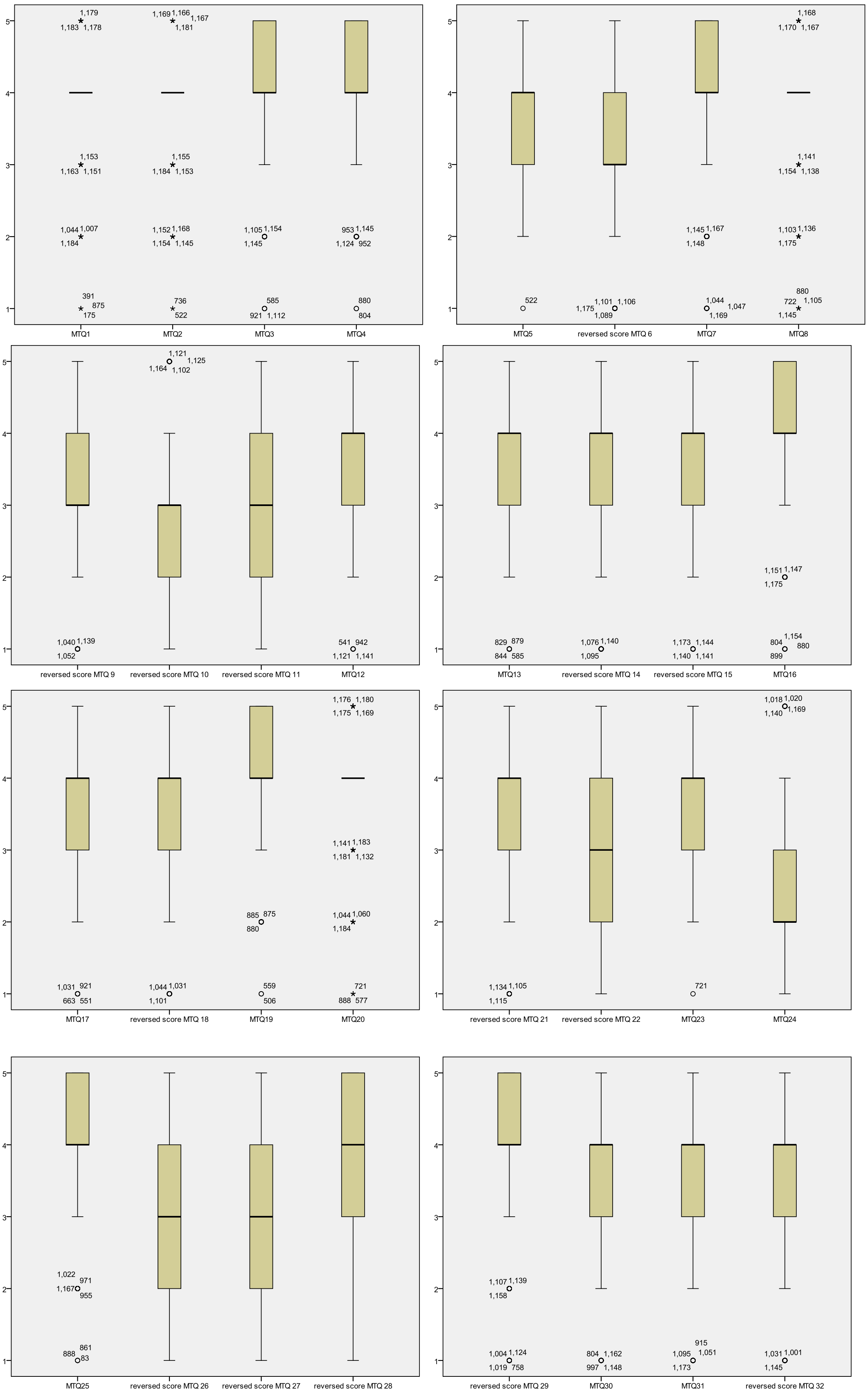
The aim of this study was to examine the suitability of the Mental Toughness Questionnaire-48 (MTQ-48) in measuring mental toughness in sport and to devise programmes to enhance the suitability of the MTQ-48.

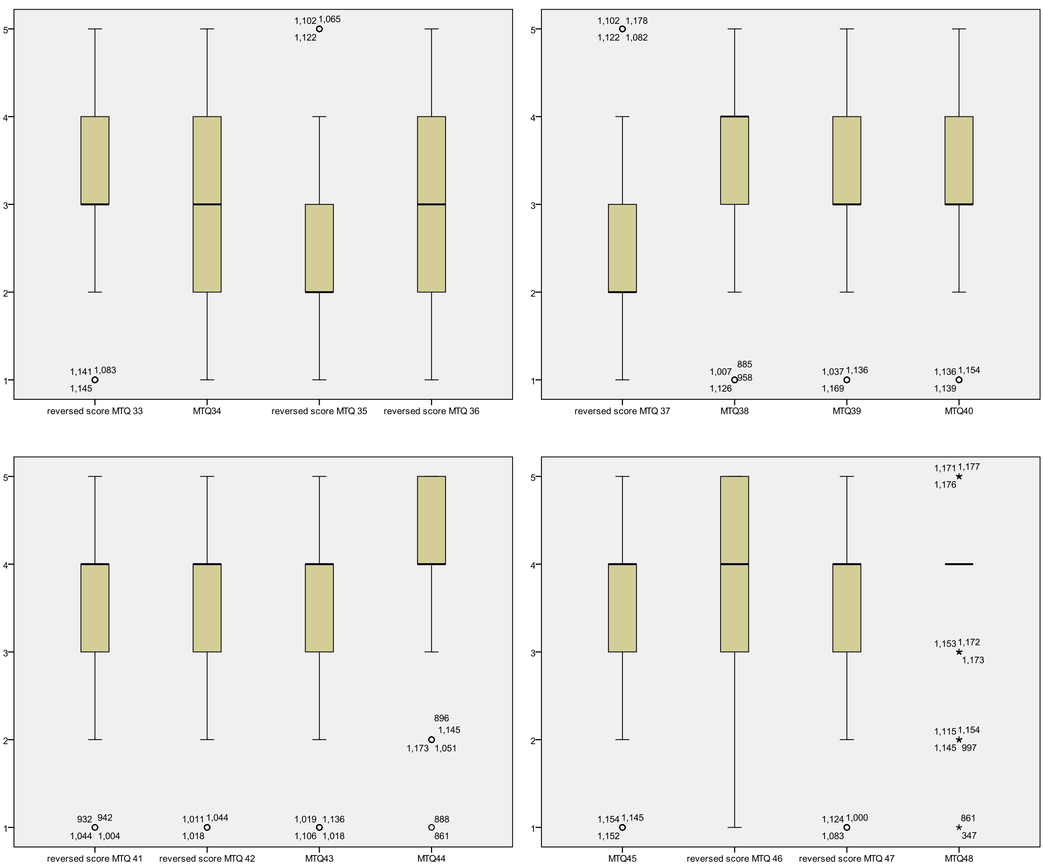
This required us to compare the MTQ-48 with another measure of mental toughness (i.e., the Sport Mental Toughness Questionnaire).

If you have any queries about this study please do not hesitate to contact Phil Birch (p.birch@chi.ac.uk) or Dr. Iain Greenlees (i.greenlees@chi.ac.uk).

Thank you again for your participation

Appendix 1.8
Individual boxplots for Mental Toughness Questionnaire-48 scores (data set A)





reversed score MTQ 33
MTQ34
reversed score MTQ 35
reversed score MTQ 36
reversed score MTQ 37
MTQ38
MTQ39
MTQ40
reversed score MTQ 41
reversed score MTQ 42
MTQ43
MTQ44
MTQ45
reversed score MTQ 46
reversed score MTQ 47
MTQ48

Appendix 1.9
Multivariate outliers (Mahalanobis distance)

Observation number	Mahalanobis d-squared	p1	p2
300	139.697	.000	.000
80	116.151	.000	.000
318	116.017	.000	.000
357	115.670	.000	.000
121	111.662	.000	.000
441	97.870	.000	.000
275	93.581	.000	.000
82	92.386	.000	.000
32	89.906	.000	.000
109	89.323	.000	.000
227	87.747	.000	.000
247	87.085	.000	.000
214	86.372	.000	.000
58	85.942	.000	.000
448	84.519	.000	.000
268	83.987	.000	.000
218	82.665	.001	.000
260	82.062	.001	.000
125	81.982	.001	.000
299	81.046	.001	.000
52	80.775	.001	.000
197	80.417	.001	.000
168	79.333	.001	.000
100	78.859	.001	.000
356	78.739	.001	.000
164	78.109	.002	.000
233	78.092	.002	.000
181	77.994	.002	.000
281	77.929	.002	.000
237	77.234	.002	.000
170	77.054	.002	.000
430	76.827	.002	.000
238	75.967	.003	.000
271	74.618	.004	.000
189	74.107	.004	.000
8	73.821	.004	.000
103	73.562	.005	.000
239	73.471	.005	.000
386	72.694	.006	.000
47	72.469	.006	.000
270	72.426	.006	.000
49	72.055	.006	.000
366	71.850	.007	.000

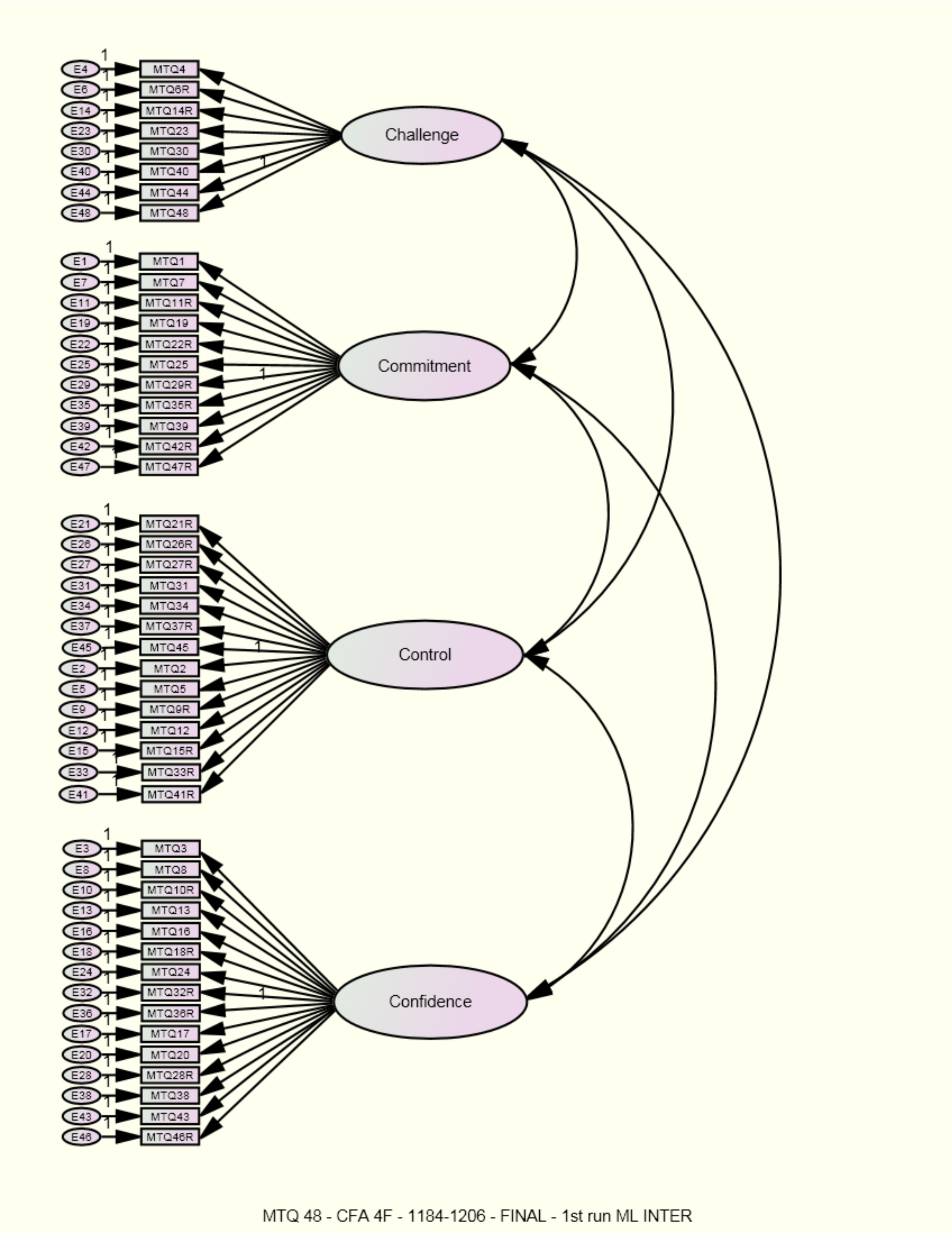
187		70.907	.008	.000
343		70.851	.008	.000
411		69.740	.010	.000
241		69.554	.011	.000
403		69.118	.012	.000
145		69.009	.012	.000
477		68.896	.012	.000
413		68.861	.013	.000
405		68.661	.013	.000
213		68.229	.014	.000
360		67.471	.017	.000
205		67.370	.017	.000
370		66.950	.018	.000
396		66.794	.019	.000
340		66.380	.021	.000
330		66.145	.022	.000
432		65.791	.023	.000
61		65.429	.025	.000
462		65.082	.027	.000
328		64.928	.027	.000
202		64.794	.028	.000
217		64.778	.028	.000
23		64.712	.029	.000
98		64.451	.030	.000
138		64.328	.031	.000
211		64.226	.031	.000
128		64.189	.032	.000
474		63.736	.034	.000
230		63.575	.035	.000
9		63.462	.036	.000
465		63.139	.038	.000
246		62.801	.041	.000
33		62.664	.042	.000
60		62.512	.043	.000
182		62.100	.046	.000
334		62.073	.046	.000
22		61.720	.049	.000
304		61.715	.049	.000
415		61.604	.050	.000
195		61.075	.055	.000
206		61.070	.055	.000
89		61.057	.056	.000
293		60.771	.058	.000
51		60.658	.060	.000
463		60.653	.060	.000
451		60.348	.063	.000
265		60.231	.064	.000

433		60.163	.065	.000
249		60.161	.065	.000
453		60.089	.066	.000
29		60.071	.066	.000
78		59.997	.067	.000
425		59.807	.069	.000
120		59.758	.069	.000
236		59.569	.072	.000
18		59.131	.077	.000
71		58.711	.082	.000

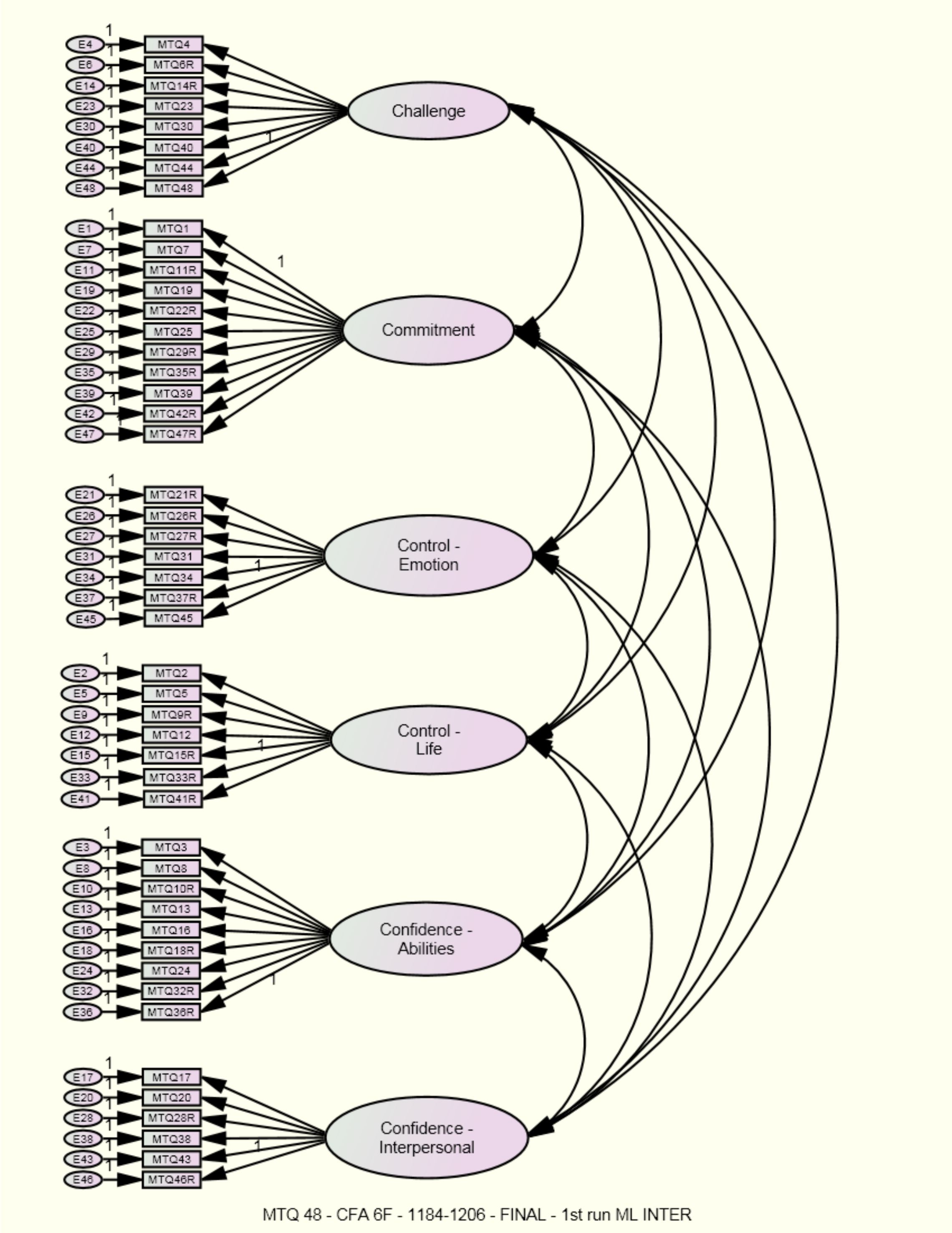
Appendix 1.10
Univariate normality statistics - extreme outliers removed (data set B)

Item	Skewness ratio	Kurtosis ratio
MTQ1	-6.65	3.18
MTQ2	-8.11	8.55
MTQ3	-9.40	8.40
MTQ4	-7.45	-1.02
MTQ5	-1.77	-2.41
MTQ6R	-5.21	-2.87
MTQ7	-13.43	7.98
MTQ8	-9.00	3.90
MTQ9R	-1.46	-2.98
MTQ10R	3.78	-3.70
MTQ11R	-1.59	-6.51
MTQ12	-9.79	5.73
MTQ13	-7.20	-0.50
MTQ14R	-6.52	-2.70
MTQ15R	-6.23	-1.59
MTQ16	-11.58	8.67
MTQ17	-6.65	-2.56
MTQ18R	-6.46	-3.14
MTQ19	-9.10	8.95
MTQ20	-7.46	1.52
MTQ21R	-8.36	-2.39
MTQ22R	-3.66	-4.93
MTQ23	-8.00	2.50
MTQ24	7.52	-1.45
MTQ25	-14.17	6.44
MTQ26R	0.64	-6.62
MTQ27R	3.60	-6.32
MTQ28R	-9.94	-0.62
MTQ29R	-16.83	14.35
MTQ30	-9.03	5.42
MTQ31	-5.14	-1.83
MTQ32R	-6.20	-1.46
RMTQ 33	-2.40	-2.26
MTQ34	-1.74	-5.95
MTQ35R	5.65	-4.37
MTQ36R	-1.85	-6.60
MTQ37R	6.83	-3.12
MTQ38	-7.34	-0.74
MTQ39	-3.03	-3.93
MTQ40	-3.70	-3.22
MTQ41R	-8.17	2.35
MTQ42R	-8.12	-0.14
MTQ43	-7.76	-1.60
MTQ44	-8.24	1.33
MTQ45	-7.19	-1.95
MTQ46R	-8.74	-1.53
MTQ47R	-7.80	-1.35
MTQ48	-5.56	6.22

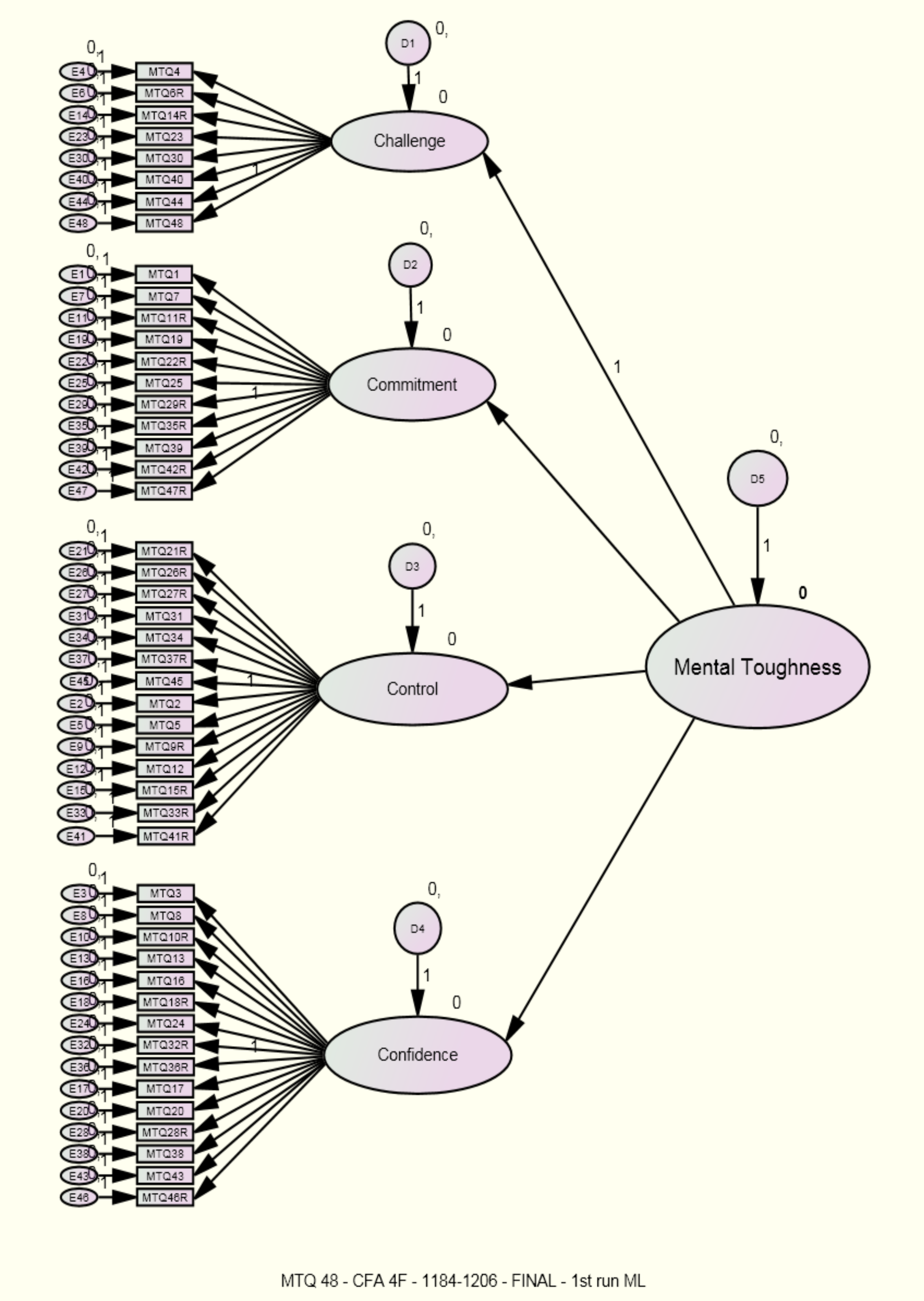
Appendix 1.11
Path diagrams of the respective specified models of the Mental Toughness Questionnaire-48



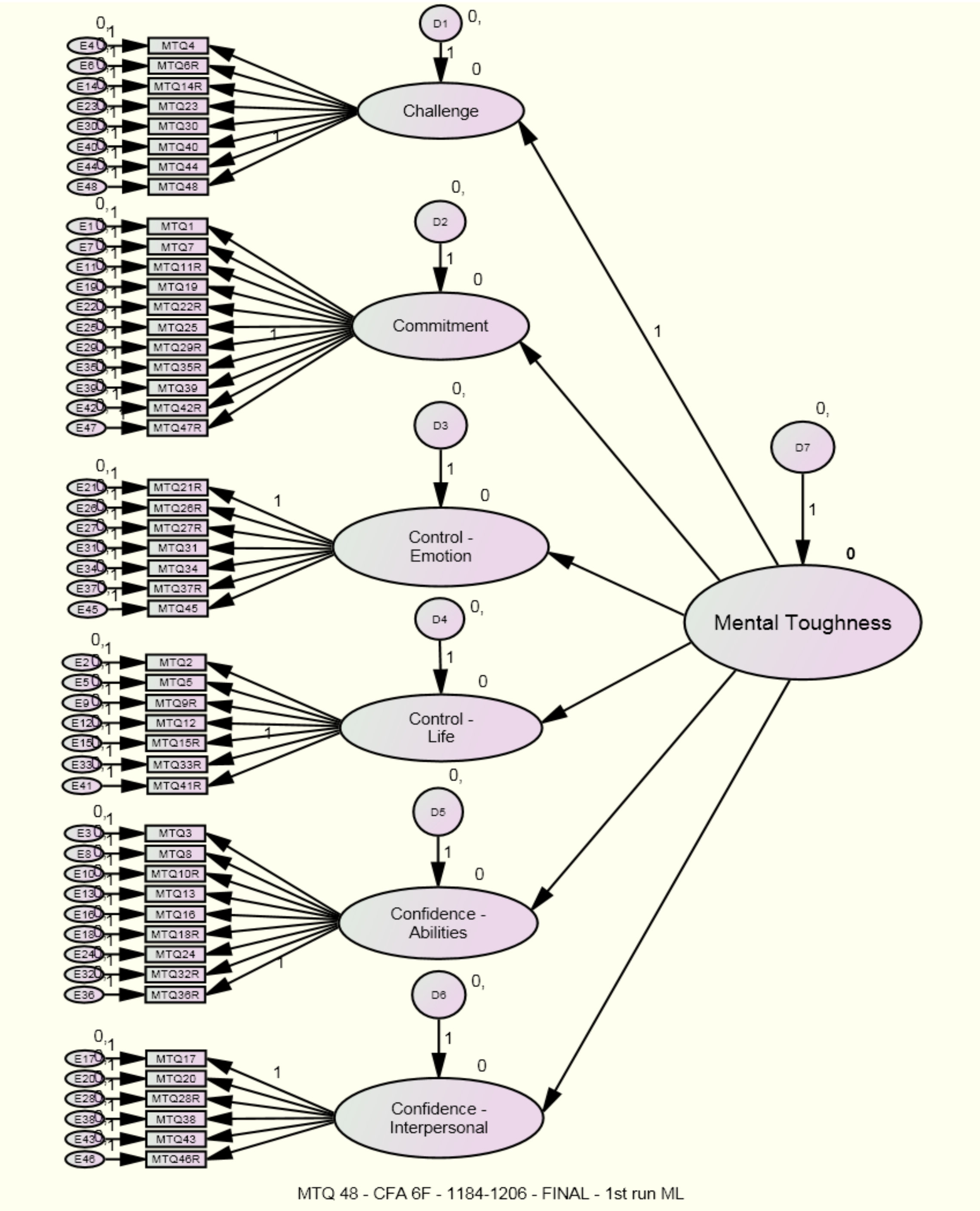
AMOS path diagram depicting first-order four factor model of mental toughness



AMOS path diagram depicting first-order six factor model of mental toughness



AMOS path diagram depicting second-order four factor model of mental toughness



AMOS path diagram depicting second-order six factor model of mental toughness

Appendix 1.12
Summary of fit indices across hypothesised model specification and estimation method of the Mental Toughness Questionnaire-48

Summary of fit indices across hypothesised model specification and estimation method of the MTQ48

	Fit indices					
CFA method	CMIN/DF	*P*	CFI	PCFI	RMSEA (90% CI)	AIC
Criterion values	(<2.00)	(> 0.05)	(>0.95)	(>0.6)	(<0.05)	(lower = better)
First-order four factor						
ML	6.129	0.000	0.614	0.585*	0.066 (0.064, 0.067)	6786.535
GLS	3.058	0.000	0.207	0.197	0.042* (0.040, 0.043)	3488.476
ADF	18.767	0.000	0.649	0.618	0.123 (0.121, 0.124)	20359.922
First-order six factor						
ML	5.334	0.000	0.677	0.639*	0.061 (0.059, 0.062)	5902.412
GLS	2.949	0.000	0.255	0.241	0.041* (0.039, 0.042)	3363.087
ADF	23.564	0.000	0.564	0.532	0.137 (0.136, 0.139)	24997.935
Second-order four factor						
ML	6.165	0.000	0.611	0.583	0.066 (0.065, 0.068)	6929.326
GLS	3.071	0.000	0.201	0.191	0.042* (0.040, 0.043)	3504.073
ADF	18.696	0.000	0.650	0.620*	0.122 (0.121, 0.124)	20317.222
Second-order six factor						
ML	5.433	0.000	0.667	0.635*	0.061 (0.060, 0.063)	6134.95
GLS	2.976	0.000	0.239	0.227	0.041* (0.039, 0.043)	3400.223
ADF	60.957	0.000	0	0	0.225 (0.224, 0.227)	65671.854

* indicates good fit.

Appendix 1.13
Regression weights of the respective models of the Mental Toughness Questionnaire-48

Standardised regression weights of the first-order four factor model

Factor / item			Estimate
MTQ44	<---	Challenge	0.614*
MTQ48	<---	Challenge	0.610*
MTQ23	<---	Challenge	0.591*
MTQ4	<---	Challenge	0.578*
MTQ30	<---	Challenge	0.496*
MTQ14R	<---	Challenge	0.389*
MTQ40	<---	Challenge	0.363*
MTQ6R	<---	Challenge	0.361*
MTQ29R	<---	Commitment	0.621*
MTQ42R	<---	Commitment	0.526*
MTQ25	<---	Commitment	0.518*
MTQ39	<---	Commitment	0.510*
MTQ7	<---	Commitment	0.504*
MTQ11R	<---	Commitment	0.480*
MTQ22R	<---	Commitment	0.462*
MTQ47R	<---	Commitment	0.454*
MTQ1	<---	Commitment	0.440*
MTQ35R	<---	Commitment	0.434*
MTQ19	<---	Commitment	0.430*
MTQ45	<---	Control	0.518*
MTQ41R	<---	Control	0.518*
MTQ2	<---	Control	0.509*
MTQ31	<---	Control	0.497*
MTQ12	<---	Control	0.477*
MTQ27R	<---	Control	0.436*
MTQ15R	<---	Control	0.424*
MTQ33R	<---	Control	0.364*
MTQ5	<---	Control	0.358*
MTQ21R	<---	Control	0.346*
MTQ37R	<---	Control	0.315*
MTQ9R	<---	Control	0.239*
MTQ26R	<---	Control	0.007
MTQ34	<---	Control	-0.136*

MTQ18R	<---	Confidence	0.556*
MTQ8	<---	Confidence	0.537*
MTQ32R	<---	Confidence	0.519*
MTQ3	<---	Confidence	0.514*
MTQ36R	<---	Confidence	0.465*
MTQ28R	<---	Confidence	0.465*
MTQ38	<---	Confidence	0.448*
MTQ16	<---	Confidence	0.437*
MTQ46R	<---	Confidence	0.434*
MTQ13	<---	Confidence	0.429*
MTQ10R	<---	Confidence	0.383*
MTQ20	<---	Confidence	0.381*
MTQ43	<---	Confidence	0.366*
MTQ17	<---	Confidence	0.322*
MTQ24	<---	Confidence	0.232*

Note: * Indicates significant relationship ($p < 0.05$).

Standardised regression weights of the first-order six factor model

Factor / item			Estimate
MTQ44	<---	Challenge	0.619*
MTQ48	<---	Challenge	0.608*
MTQ23	<---	Challenge	0.592*
MTQ4	<---	Challenge	0.577*
MTQ30	<---	Challenge	0.501*
MTQ14R	<---	Challenge	0.380*
MTQ40	<---	Challenge	0.361*
MTQ6R	<---	Challenge	0.360*
MTQ29R	<---	Commitment	0.618*
MTQ42R	<---	Commitment	0.531*
MTQ25	<---	Commitment	0.520*
MTQ39	<---	Commitment	0.508*
MTQ7	<---	Commitment	0.499*
MTQ47R	<---	Commitment	0.458*
MTQ11R	<---	Commitment	0.477*
MTQ22R	<---	Commitment	0.462*
MTQ1	<---	Commitment	0.445*
MTQ19	<---	Commitment	0.439*
MTQ35R	<---	Commitment	0.427*
MTQ45	<---	Control - Emotion	0.628*
MTQ31	<---	Control - Emotion	0.616*
MTQ27R	<---	Control - Emotion	0.549*
MTQ21R	<---	Control - Emotion	0.406*
MTQ37R	<---	Control - Emotion	0.297*
MTQ26R	<---	Control - Emotion	0.077
MTQ34	<---	Control - Emotion	-0.087
MTQ41R	<---	Control - Life	0.558*
MTQ2	<---	Control - Life	0.513*
MTQ12	<---	Control - Life	0.498*
MTQ15R	<---	Control - Life	0.449*
MTQ33R	<---	Control - Life	0.403*
MTQ5	<---	Control - Life	0.343*
MTQ9R	<---	Control - Life	0.248*

MTQ18R	<---	Confidence - Abilities	0.577*
MTQ32R	<---	Confidence - Abilities	0.560*
MTQ3	<---	Confidence - Abilities	0.540*
MTQ8	<---	Confidence - Abilities	0.532*
MTQ16	<---	Confidence - Abilities	0.501*
MTQ13	<---	Confidence - Abilities	0.482*
MTQ36R	<---	Confidence - Abilities	0.480*
MTQ10R	<---	Confidence - Abilities	0.409*
MTQ24	<---	Confidence - Abilities	0.278*
MTQ43	<---	Confidence -Interpersonal	0.643*
MTQ38	<---	Confidence - Interpersonal	0.639*
MTQ17	<---	Confidence -Interpersonal	0.580*
MTQ20	<---	Confidence -Interpersonal	0.578*
MTQ46R	<---	Confidence -Interpersonal	0.572*
MTQ28R	<---	Confidence -Interpersonal	0.460*

Note: * Indicates significant relationship ($p < 0.05$).

Standardised regression weights of the second-order four factor model

Factor / item			Estimate
Control	<---	Mental Toughness	1.019*
Confidence	<---	Mental Toughness	0.953*
Challenge	<---	Mental Toughness	0.851*
Commitment	<---	Mental Toughness	0.744*
MTQ48	<---	Challenge	0.603*
MTQ23	<---	Challenge	0.598*
MTQ44	<---	Challenge	0.597*
MTQ4	<---	Challenge	0.571*
MTQ30	<---	Challenge	0.507*
MTQ14R	<---	Challenge	0.392*
MTQ40	<---	Challenge	0.373*
MTQ6R	<---	Challenge	0.373*
MTQ29R	<---	Commitment	0.626*
MTQ42R	<---	Commitment	0.531*
MTQ25	<---	Commitment	0.509*
MTQ39	<---	Commitment	0.507*
MTQ7	<---	Commitment	0.497*
MTQ11R	<---	Commitment	0.487*
MTQ22R	<---	Commitment	0.466*
MTQ47R	<---	Commitment	0.457*
MTQ35R	<---	Commitment	0.440*
MTQ1	<---	Commitment	0.437*
MTQ19	<---	Commitment	0.421*
MTQ45	<---	Control	0.521*
MTQ2	<---	Control	0.515*
MTQ41R	<---	Control	0.511*
MTQ31	<---	Control	0.511*
MTQ12	<---	Control	0.477*
MTQ27R	<---	Control	0.431*
MTQ15R	<---	Control	0.421*
MTQ33R	<---	Control	0.357*
MTQ5	<---	Control	0.356*
MTQ21R	<---	Control	0.341*

MTQ37R	<---	Control	0.318*
MTQ9R	<---	Control	0.240*
MTQ26R	<---	Control	0.012
MTQ34	<---	Control	-0.127*
MTQ18R	<---	Confidence	0.556*
MTQ8	<---	Confidence	0.541*
MTQ32R	<---	Confidence	0.517*
MTQ3	<---	Confidence	0.515*
MTQ28R	<---	Confidence	0.464*
MTQ36R	<---	Confidence	0.456*
MTQ38	<---	Confidence	0.449*
MTQ16	<---	Confidence	0.438*
MTQ46R	<---	Confidence	0.437*
MTQ13	<---	Confidence	0.428*
MTQ20	<---	Confidence	0.386*
MTQ10R	<---	Confidence	0.384*
MTQ43	<---	Confidence	0.366*
MTQ17	<---	Confidence	0.324*
MTQ24	<---	Confidence	0.226*

Note: * Indicates significant relationship (p < 0.05).

Standardised regression weights of the second-order six factor model

Factor / item			Estimate
Control - Life	<---	Mental Toughness	0.958*
Confidence -Abilities	<---	Mental Toughness	0.902*
Challenge	<---	Mental Toughness	0.875*
Control - Emotion	<---	Mental Toughness	0.798*
Commitment	<---	Mental Toughness	0.765*
Confidence - Interpersonal	<---	Mental Toughness	0.574*
MTQ48	<---	Challenge	0.606*
MTQ44	<---	Challenge	0.599*
MTQ23	<---	Challenge	0.598*
MTQ4	<---	Challenge	0.573*
MTQ30	<---	Challenge	0.501*
MTQ14R	<---	Challenge	0.395*
MTQ40	<---	Challenge	0.371*
MTQ6R	<---	Challenge	0.367*
MTQ29R	<---	Commitment	0.624*
MTQ42R	<---	Commitment	0.531*
MTQ25	<---	Commitment	0.513*
MTQ39	<---	Commitment	0.507*
MTQ7	<---	Commitment	0.500*
MTQ11R	<---	Commitment	0.484*
MTQ22R	<---	Commitment	0.464*
MTQ47R	<---	Commitment	0.458*
MTQ1	<---	Commitment	0.440*
MTQ35R	<---	Commitment	0.433*
MTQ19	<---	Commitment	0.426*
MTQ41R	<---	Control - Life	0.539*
MTQ2	<---	Control - Life	0.537*
MTQ12	<---	Control - Life	0.503*
MTQ15R	<---	Control - Life	0.445*
MTQ33R	<---	Control - Life	0.390*
MTQ5	<---	Control - Life	0.335*

MTQ9R	<---	Control - Life	0.252*
MTQ31	<---	Control - Emotion	0.612*
MTQ45	<---	Control - Emotion	0.628*
MTQ27R	<---	Control - Emotion	0.543*
MTQ21R	<---	Control - Emotion	0.398*
MTQ37R	<---	Control - Emotion	0.323*
MTQ26R	<---	Control - Emotion	0.078
MTQ34	<---	Control - Emotion	-0.088
MTQ18R	<---	Confidence - Abilities	0.576*
MTQ3	<---	Confidence - Abilities	0.551*
MTQ8	<---	Confidence - Abilities	0.549*
MTQ32R	<---	Confidence - Abilities	0.543*
MTQ16	<---	Confidence - Abilities	0.513*
MTQ13	<---	Confidence - Abilities	0.488*
MTQ36R	<---	Confidence - Abilities	0.457*
MTQ10R	<---	Confidence - Abilities	0.406*
MTQ24	<---	Confidence - Abilities	0.272*
MTQ43	<---	Confidence - Interpersonal	0.642*
MTQ38	<---	Confidence - Interpersonal	0.639*
MTQ17	<---	Confidence - Interpersonal	0.579*
MTQ46R	<---	Confidence - Interpersonal	0.575*
MTQ20	<---	Confidence - Interpersonal	0.569*
MTQ28R	<---	Confidence - Interpersonal	0.470*

Note: * Indicates significant relationship ($p < 0.05$).

Appendix 1.14
Independent model re-specifications of the respective models of the Mental Toughness Questionnaire

Fit indices of the first-order four factor model re-specification

		Fit indices					
CFA run Criterion values	Items removed	CMIN/DF (< 2.00)	*P* (> 0.05)	CFI (> 0.95)	PCFI (> 0.6)	RMSEA (90% CI) (< 0.05)	AIC (lower = better)
1	Hypothesised model	6.129	0.000	0.614	0.585	0.066 (0.064, 0.067)	6786.535
2	34, 26R	6.114	0.000	0.635	0.603*	0.066 (0.064, 0.067)	6206.216
3	13	6.032	0.000	0.644	0.611*	0.065 (0.064, 0.067)	5856.435
4	17	5.908	0.000	0.657	0.623*	0.064 (0.063, 0.066)	5481.366
5	20	5.762	0.000	0.671	0.635*	0.063 (0.062, 0.065)	5104.643
6	43	5.617	0.000	0.686	0.647*	0.062 (0.061, 0.064)	4746.276
7	33R	5.551	0.000	0.695	0.655*	0.062 (0.060, 0.064)	4467.123
8	46R	5.590	0.000	0.700	0.658*	0.062 (0.060, 0.064)	4275.355
9	35R	5.275	0.000	0.720	0.676*	0.060 (0.058, 0.062)	3839.625
10	14R, 38, 6R, 40, 21R, 5, 37R, 24, 9R	5.480	0.000	0.788	0.723*	0.062 (0.059, 0.064)	2318.705

* indicates good fit.

Fit indices of the first-order six factor model re-specification

		Fit indices					
CFA run Criterion values	Items removed	CMIN/DF (< 2.00)	*P* (> 0.05)	CFI (> 0.95)	PCFI (> 0.6)	RMSEA (90% CI) (< 0.05)	AIC (lower = better)
1	Hypothesised model	5.334	0.000	0.677	0.639*	0.061 (0.059, 0.062)	5902.412
2	34, 26R	5.238	0.000	0.700	0.659*	0.060 (0.058, 0.061)	5316.17
3	13	5.163	0.000	0.708	0.665*	0.059 (0.058, 0.061)	5011.692
4	33R	5.101	0.000	0.717	0.673*	0.059 (0.057, 0.061)	4730.473
5	35R	4.834	0.000	0.736	0.689*	0.057 (0.055, 0.059)	4286.854
6	4[a]	4.924	0.000	0.733	0.685*	0.058 (0.056, 0.059)	4157.16
7	36R	4.767	0.000	0.744	0.695*	0.056 (0.055, 0.058)	4030.514
8	22R[a]	4.815	0.000	0.747	0.696*	0.057 (0.055, 0.059)	3872.884
9	6R	4.656	0.000	0.757	0.705*	0.056 (0.054, 0.057)	3751.455
10	21R, 5, 14R, 40, 37R, 24, 9R	4.723	0.000	0.798	0.731*	0.056 (0.054, 0.058)	2744.062

* indicates good fit.

[a] denotes item being retained due to decrement in model fit

Fit indices of the second-order four factor model re-specification

CFA run Criterion values	Items removed	Fit indices					
		CMIN/DF (< 2.00)	*P* (> 0.05)	CFI (> 0.95)	PCFI (> 0.6)	RMSEA (90% CI) (< 0.05)	AIC (lower = better)
1	Hypothesised model	6.165	0.000	0.611	0.583	0.066 (0.065, 0.068)	6929.326
2	34	6.070	0.000	0.625	0.596	0.065 (0.064, 0.067)	6542.103
3	26R	6.151	0.000	0.632	0.601*	0.066 (0.064, 0.068)	6342.637
4	13	6.069	0.000	0.641	0.609*	0.065 (0.064, 0.067)	5988.914
5	17	5.948	0.000	0.654	0.621*	0.065 (0.063, 0.066)	5613.691
6	20	5.812	0.000	0.667	0.632*	0.064 (0.062, 0.066)	5240.997
7	43	5.671	0.000	0.681	0.645*	0.063 (0.061, 0.065)	4882.063
8	33R	5.600	0.000	0.691	0.653*	0.062 (0.061, 0.064)	4594.223
9	14R, 6R, 40, 21R, 5, 37R, 9R, 24	5.540	0.000	0.760	0.707*	0.062 (0.060, 0.064)	2926.319

* indicates good fit.

Fit indices of the second-order six factor model re-specification

CFA run Criterion values	Items removed	CMIN/DF (< 2.00)	*P* (> 0.05)	CFI (> 0.95)	PCFI (> 0.6)	RMSEA (90% CI) (< 0.05)	AIC (lower = better)
		Fit indices					
1	Hypothesised model	5.433	0.000	0.667	0.635*	0.061 (0.060, 0.063)	6134.95
2	36<>26	5.179	0.000	0.686	0.653*	0.059 (0.058, 0.061)	5858.948
3	34	5.295	0.000	0.683	0.650*	0.060 (0.059, 0.062)	5737.558
4	26R	5.348	0.000	0.690	0.655*	0.061 (0.059, 0.062)	5545.051
5	13	5.280	0.000	0.697	0.661*	0.060 (0.059, 0.062)	5240.085
6	33R	5.205	0.000	0.707	0.670*	0.060 (0.058, 0.061)	4939.714
7	4[a]	5.304	0.000	0.704	0.666*	0.060 (0.059, 0.062)	4799.538
8	35R	4.954	0.000	0.725	0.686*	0.058 (0.056, 0.060)	4500.99
9	17[a]	5.026	0.000	0.724	0.683*	0.058 (0.057, 0.060)	4349.989
10	22R	5.021	0.000	0.727	0.686*	0.058 (0.057, 0.060)	4346.154
11	14R, 21R, 40, 6R, 5, 37R, 24, 9R	5.155	0.000	0.783	0.727*	0.059 (0.057, 0.061)	2901.665

* indicates good fit.

[a] denotes item being retained due to decrement in model fit

Study 2: Assessment Instruments and SPSS Outputs

Appendix 2.1
The University of Chichester Mental Toughness Questionnaire

Considering how you are **generally**, please answer the following statements. Start each statement with the phrase "In general." For example, the first statement should be read "In general, I am confident in my own abilities." Please answer **honestly**, we are looking for your initial response to each statement so try not to spend too much time on any one statement.

Please indicate your response by **circling one** of the numbers, which have the following meaning;

1 = strongly disagree; **2** = disagree; **3** = neither agree nor disagree; **4** = agree; **5** = strongly agree.

In general, …	Strongly disagree	Disagree	Neither agree or disagree	Agree	Strongly agree
1. I am confident in my own abilities	1	2	3	4	5
2. I am in control of my life	1	2	3	4	5
3. I find that challenges stop me reaching my goals	1	2	3	4	5
4. I am able to control the impact that my nerves have on me	1	2	3	4	5
5. I doubt myself when I have a difficult task to achieve	1	2	3	4	5
6. I feel that people don't listen to what I have to say	1	2	3	4	5
7. For me, an activity is only worthwhile doing if it stretches me to my limits	1	2	3	4	5
8. If something is worth doing, I devote all my efforts to see it through	1	2	3	4	5
9. I believe that whatever is going to happen is going to happen	1	2	3	4	5
10. I find it difficult to control my emotions in high pressure situations	1	2	3	4	5
11. I see testing situations as opportunities for me to develop as a person	1	2	3	4	5
12. I let my anxieties get the better of me	1	2	3	4	5
13. I remain committed to my goals no matter what obstacles are put in front of me	1	2	3	4	5
14. I am the major cause of my own destiny	1	2	3	4	5
15. I doubt myself when faced with setbacks	1	2	3	4	5
16. My emotions get the better of me when I want to do well	1	2	3	4	5
17. I believe that I am in control of the plans I make	1	2	3	4	5
18. I find it easy to meet new people	1	2	3	4	5

In general, …	Strongly disagree	Disagree	Neither agree or disagree	Agree	Strongly agree
19. My self-belief enables me to achieve my goals	1	2	3	4	5
20. I find it difficult to stay dedicated to a task when I have a tough deadline	1	2	3	4	5
21. I believe that what will be will be	1	2	3	4	5
22. I am an assertive person	1	2	3	4	5
23. I often hide my emotions from others	1	2	3	4	5
24. I expect to succeed in performing important tasks	1	2	3	4	5
25. I often find myself giving up on a task	1	2	3	4	5
26. I think that most people I know like me	1	2	3	4	5
27. I thrive on pushing myself to the limits of my abilities	1	2	3	4	5
28. When things go wrong my head can drop	1	2	3	4	5
29. I feel that I have little control over the direction my life is taking	1	2	3	4	5
30. I am fully committed to achieving the goals I have set myself	1	2	3	4	5
31. I believe that getting what you want from life has little to do with luck	1	2	3	4	5
32. I get most enjoyment out of putting myself in challenging situations	1	2	3	4	5
33. I get frustrated when things do not go my way	1	2	3	4	5
34. I believe that what happens to me is down to my own actions	1	2	3	4	5
35. I am unable to bounce back following failures	1	2	3	4	5
36. I lack self-belief	1	2	3	4	5
37. I am a good person to be around	1	2	3	4	5
38. I find myself giving up on things when the going gets really tough	1	2	3	4	5
39. I find it difficult to manage the many competing demands in my life	1	2	3	4	5
40. I thrive in continually changing environments	1	2	3	4	5
41. No matter how hard things get, I see a task through to the end	1	2	3	4	5
42. I get intimidated in social situations	1	2	3	4	5
43. I can remain calm even in the most difficult situations	1	2	3	4	5
44. I make myself do things whether I want to or not	1	2	3	4	5
45. I am not able to deal with awkward people	1	2	3	4	5

Appendix 2.2
The University of Chichester Mental Toughness Questionnaire Scoring Key

Constructs	Items	Total number of items
Challenge	3R, 7, 11, 27, 32, 40.	6
Commitment	8, 13, 20R, 25R, 30, 35R, 38R, 41, 44.	9
Control - emotion	4, 10R, 12R, 16R, 23, 33R, 43.	7
Control - life	2, 9R, 14, 17, 21R, 31, 34, 39R.	8
Confidence – abilities	1, 5R, 15R, 19, 24, 28R, 36R.	7
Confidence - interpersonal	6R, 18, 22, 26, 29R, 37, 42R, 45R.	8

Calculating scores

Total mean mental toughness = Add all 45 items scores (accounting for reversed items!) and divide by 45.

Mean construct (subscale) scores = Add the respective items scores (accounting for reversed items!) and divide by the amount of items for that construct.

Appendix 2.3
Participant information sheet, consent form, and demographic questionnaire

Department of Sport and Exercise Sciences
Project information sheet for participants

PLEASE READ THE FOLLOWING CAREFULLY
Study title: An examination of the psychological characteristics and thought processes of competitive athletes.

What is the purpose of the research and how will the research be carried out?
The purpose of this study is to examine the psychological characteristics and thought processes of competitive athletes.

At the University of Chichester before any project starts it has to be checked by University staff. They make sure that the research is good enough to carry out. This project has been checked by the staff of the Department of Sport and Exercise Science and has been passed by the University Ethics Committee.

What will you be asked to do?
The study involves completing two questionnaires which will assess your psychological characteristics and thought processes. Please take your time to complete the questionnaires and answer as truthfully as possible.

In total, each questionnaire should take no longer than 6 minutes so the total amount of time you will need to commit to this project is 12 minutes.

What are the anticipated benefits of participating in the research?
You have been asked to take part in this project since we need athletes who are competing in sport to gain an insight into the psychology of sport. It is hoped that the research will enhance our knowledge of what attributes contribute to sporting success.

Are there any risks associated with participating in the research?
You will be providing potentially sensitive information about perceptions of your psychological characteristics and thought processes. As such, while the information you provide will be used in academic publications, your name will not be mentioned in any literature, and as such, the responses you make would not be linked back to you.

Do you have to take part?
No! It is your choice whether you participate in the project or not. If at any time during the project you feel that you don't want to continue then you can tell the researcher that you want to stop – you do not have to give a reason.

What will happen to the information collected as part of the study?
All the information which is collected about you during the project will be kept private and stored in lockable cabinets – no one will know that the information belongs to you and all the information will be kept at the University of Chichester. Your responses and information will be treated confidentially. Questionnaire responses will be associated with an anonymous code rather than a person's name. The questionnaires will be stored for a period of 7 years and then destroyed. If you wish to withdraw your data from the study you should do so within 6 months of completing the study.

Who can you contact if you have a complaint about the project?
Dr Andy Dixon – Head of the Research and Employer Engagement Office – Email a.dixon@chi.ac.uk: Phone 01243 812125

Who can you contact if you have any questions about the project?
Phil Birch – Email p.birch@chi.ac.uk: Phone 01243 816345 or
Dr Iain Greenlees - Email i.greenlees@chi.ac.uk: Phone 01243 816437

Version 1, 28/10/2011 **Thank you for your time**

Participant consent form

PLEASE READ THE FOLLOWING CAREFULLY

Study title: An examination of the psychological characteristics and thought processes of competitive athletes.

I understand that my participation in this project will involve completing two questionnaires which will assess my perceptions of psychological characteristics and thought processes.

I am happy to participate in this research Yes No

1) I have read and understand the information sheet for this research project. I have had the opportunity to consider the information, ask questions and have had these answered satisfactorily. Yes No

2) I understand that my participation in the activity is voluntary and that I am therefore free to withdraw my involvement at any stage, without giving a reason. Yes No

3) I understand that all information will be anonymised and that my personal information will not be released to any third parties. Yes No

If you responded "Yes" to the above statements please complete the following:

Your name (please print)………………………………………………….

Your signature…………………………………………………………………….

Date...

Thank you for your time

Demographic Questionnaire

Please fill in your details in the spaces provided

Age ______________________________

(years)

Gender Male ☐ Female ☐

Ethnic origin – Please indicate the most appropriate description of your ethnic origin.

(Please tick)

White - British	☐	Asian - Indian	☐
White - Irish	☐	Asian - Pakistani	☐
White – Any other White background	☐	Asian – Bangladeshi	☐
Mixed – White & Black Caribbean	☐	Black or Black British - Caribbean	☐
Mixed – White & Black African	☐	Black or Black British - African	☐
Mixed – White & Black Asian	☐	Black or Black British - Any other Black background	☐
Mixed – Any other Mixed background	☐	Chinese – Any Chinese background	☐
Any other ethnic background	☐		

University of study ____

Year of study ____

(e.g., 1, 2, 3)

What is your primary sport? ____

How many years competitive playing experience do you have in your primary sport? ___________________________

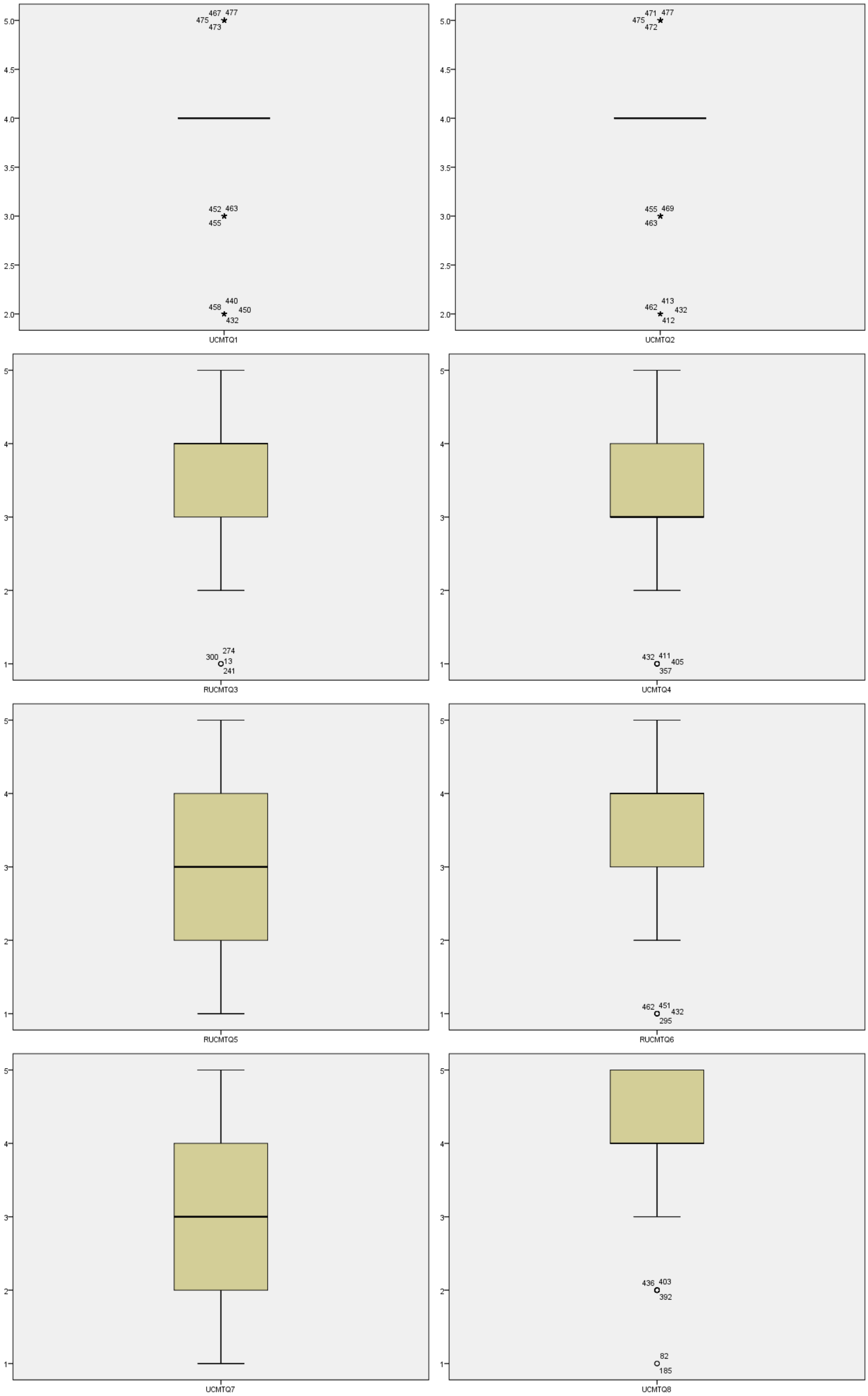
What is the highest level that you have played in your primary sport?
____

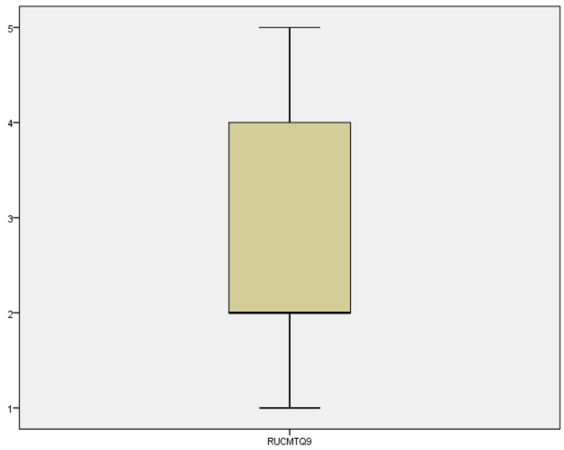
Participant debrief

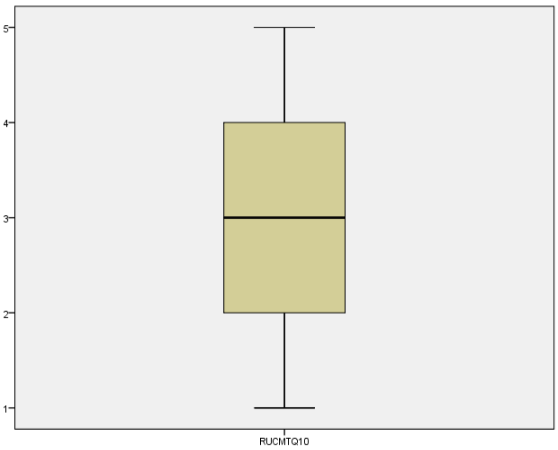
The aim of this study was to examine the suitability of University of Chichester Mental Toughness Questionnaire (UCMTQ) in measuring mental toughness in sport and to devise programmes to enhance the suitability of the UCMTQ.
This required us to compare the UCMTQ with a Social Desirability Scale.
If you have any queries about this study please do not hesitate to contact Phil Birch (p.birch@chi.ac.uk) or Dr. Iain Greenlees (i.greenlees@chi.ac.uk).

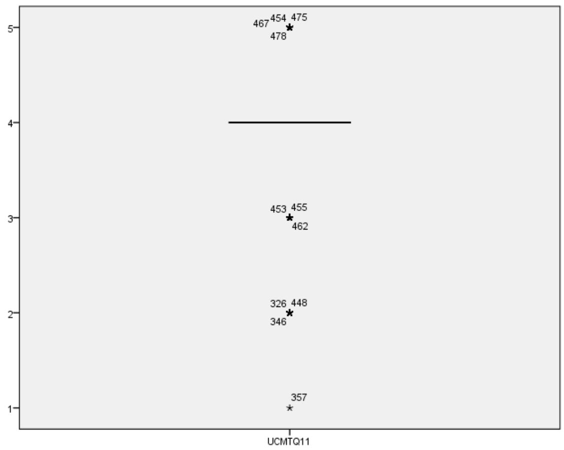
Thank you again for your participation

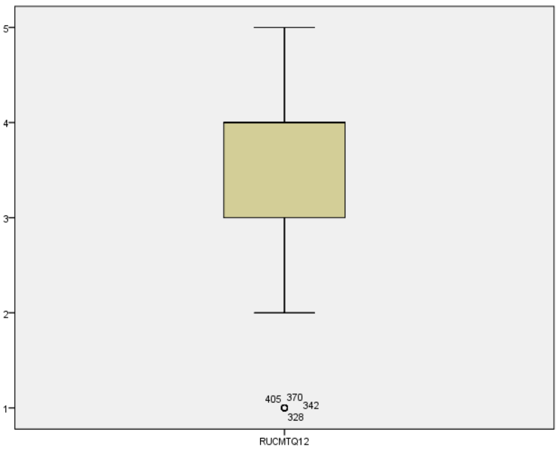
Appendix 2.4
Individual boxplots for UCMTQ scores (data set A)

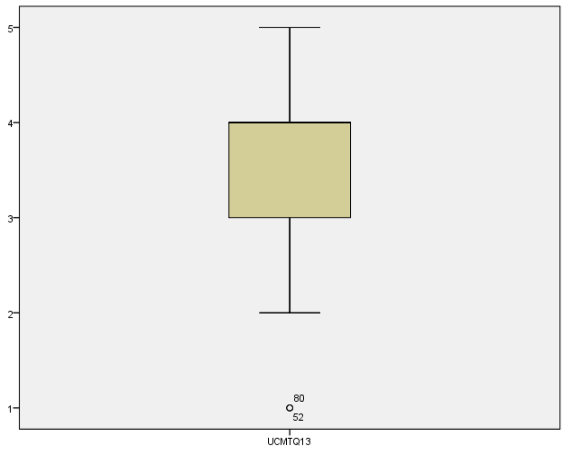


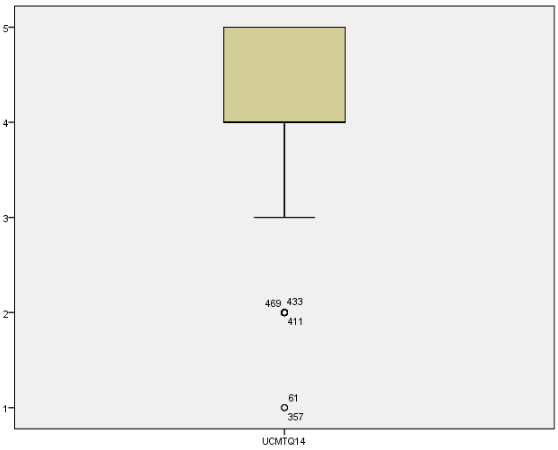

5
4
3
2
1
RUCMTQ9

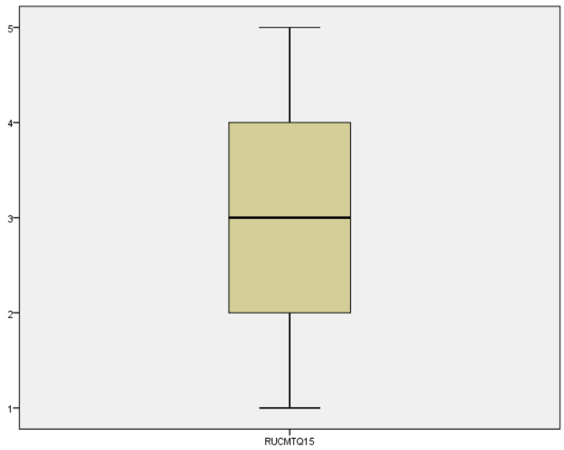

5
4
3
2
1
RUCMTQ10

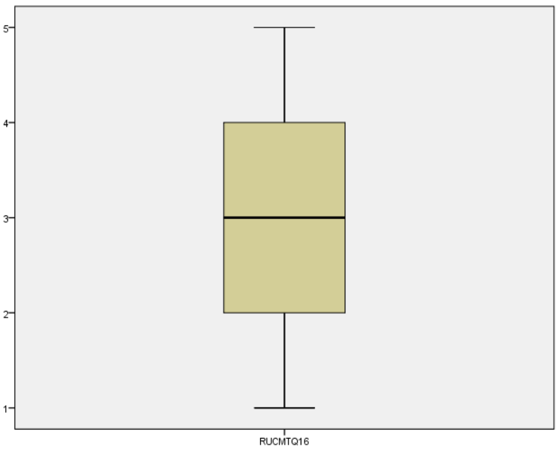

467 454 475
478
453 455
462
326 448
346
357
5
4
3
2
1
UCMTQ11

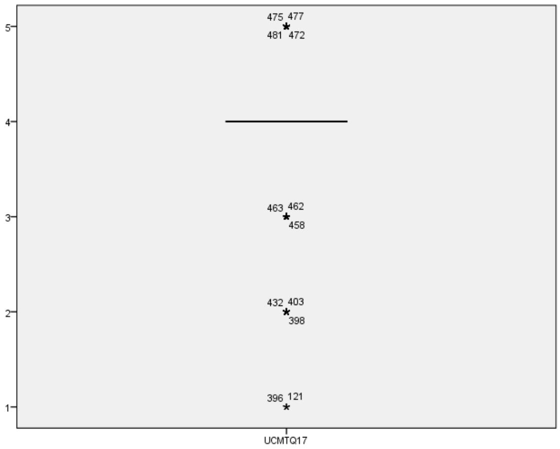

405 370
342
328
5
4
3
2
1
RUCMTQ12

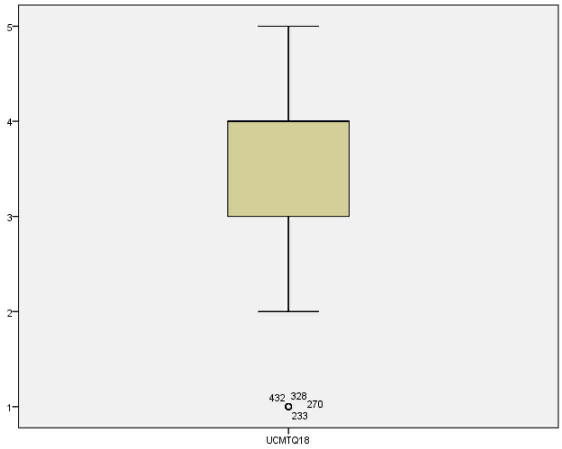

80
52
5
4
3
2
1
UCMTQ13

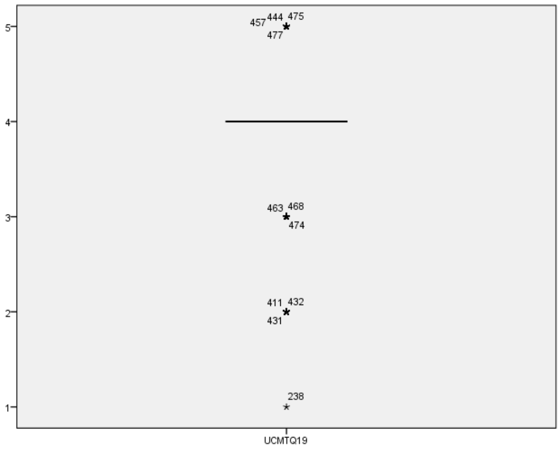

469 433
411
61
357
5
4
3
2
1
UCMTQ14

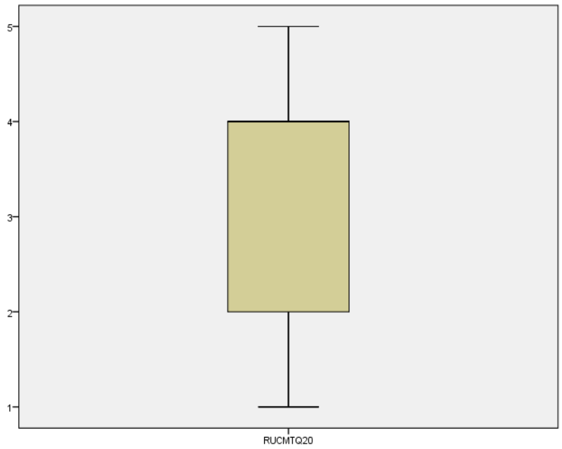

5
4
3
2
1
RUCMTQ15

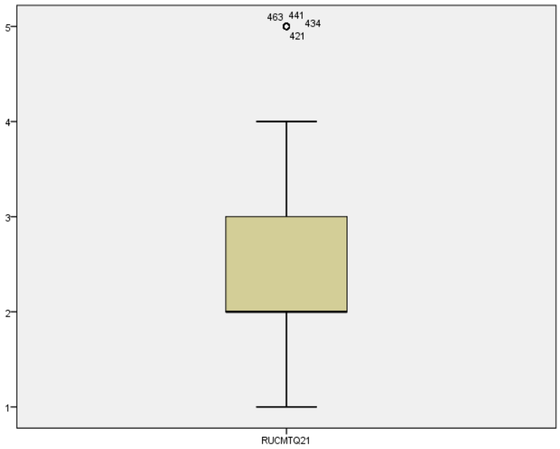

5
4
3
2
1
RUCMTQ16

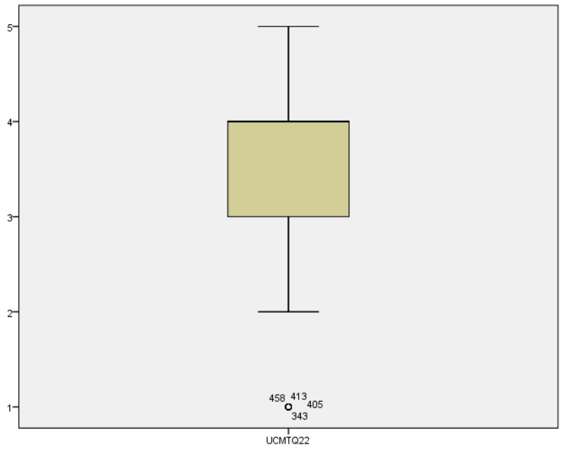

475 477
481 472
463 462
458
432 403
398
396 121
1
2
3
4
5
UCMTQ17

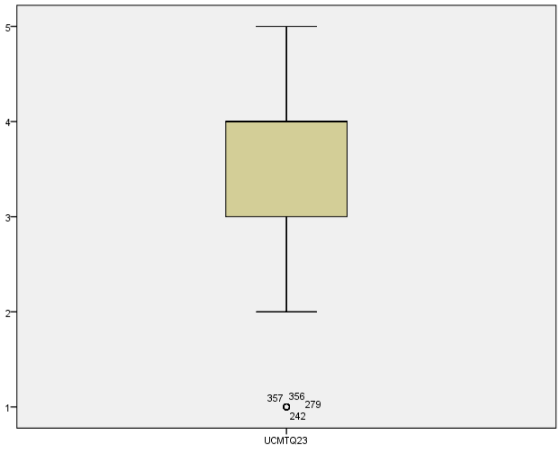

432 328
270
233
1
2
3
4
5
UCMTQ18

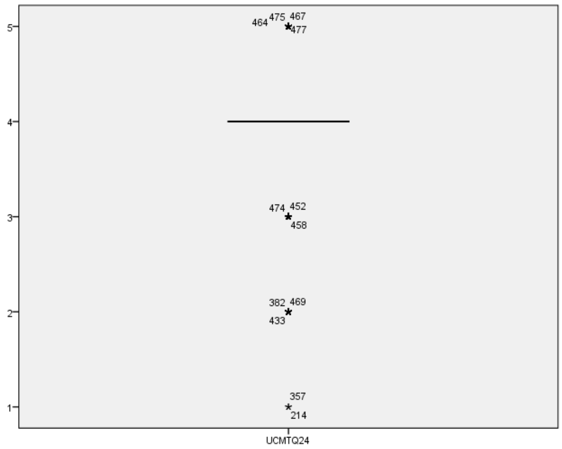

457 444 475
477
463 468
474
411 432
431
238
1
2
3
4
5
UCMTQ19

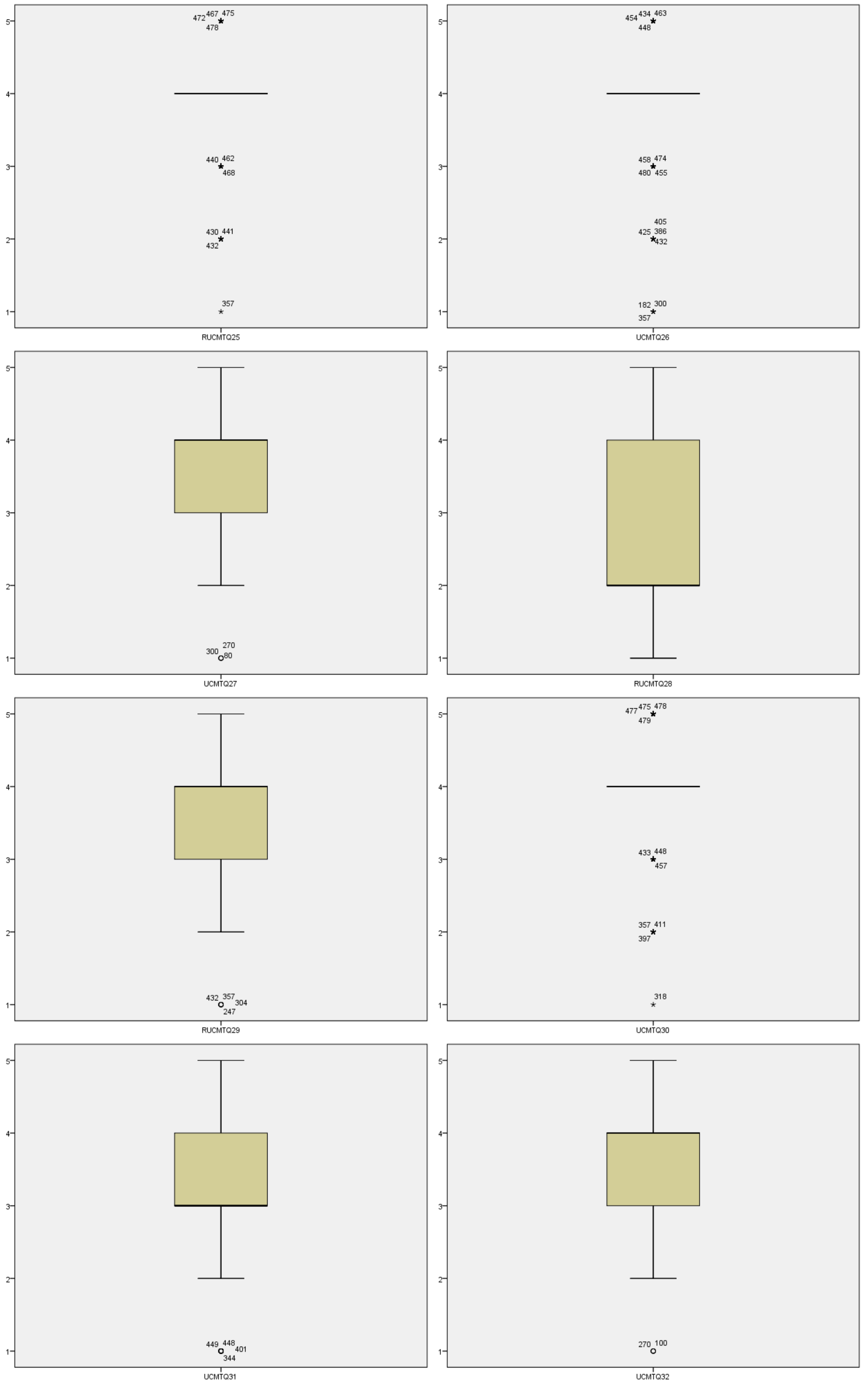

1
2
3
4
5
RUCMTQ20

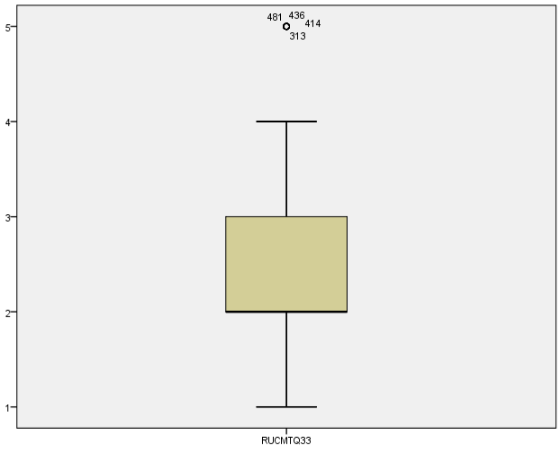

463 441
434
421
1
2
3
4
5
RUCMTQ21

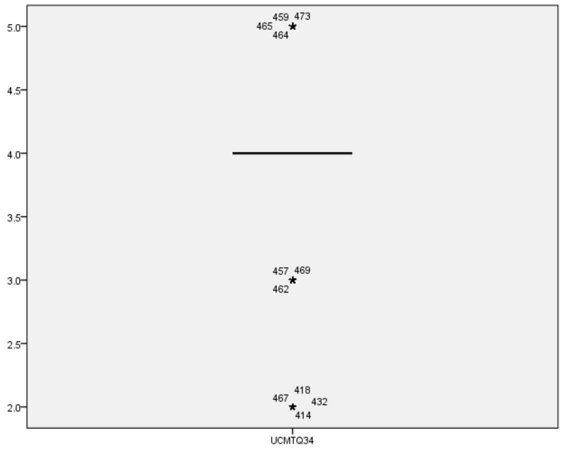

458 413
405
343
1
2
3
4
5
UCMTQ22

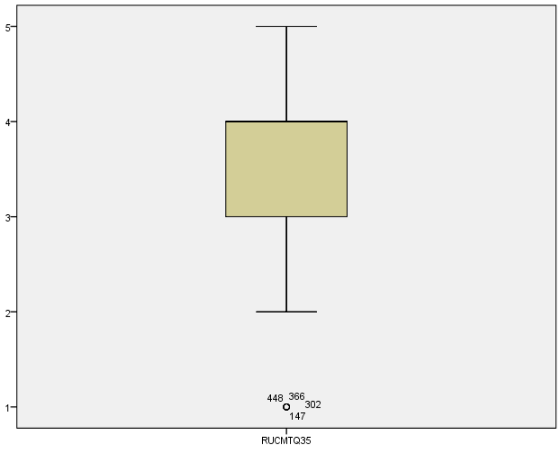

357 356
279
242
1
2
3
4
5
UCMTQ23

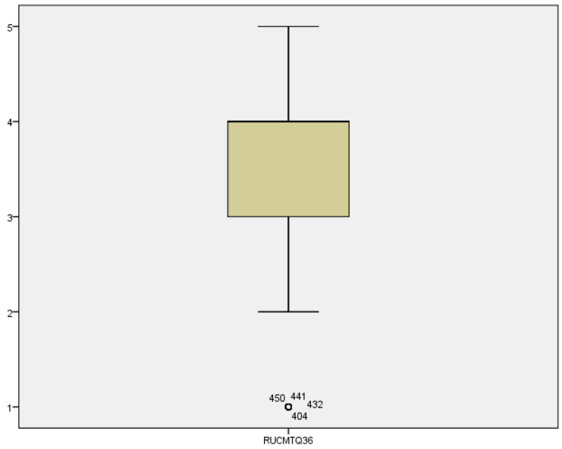

464 475 467
477
474 452
458
382 469
433
357
214
1
2
3
4
5
UCMTQ24

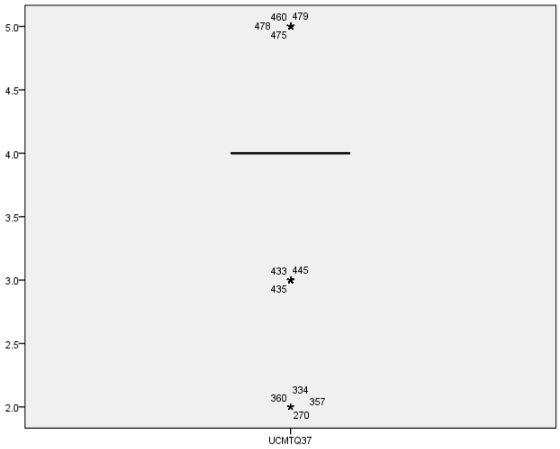

RUCMTQ25
UCMTQ26
UCMTQ27
RUCMTQ28
RUCMTQ29
UCMTQ30
UCMTQ31
UCMTQ32

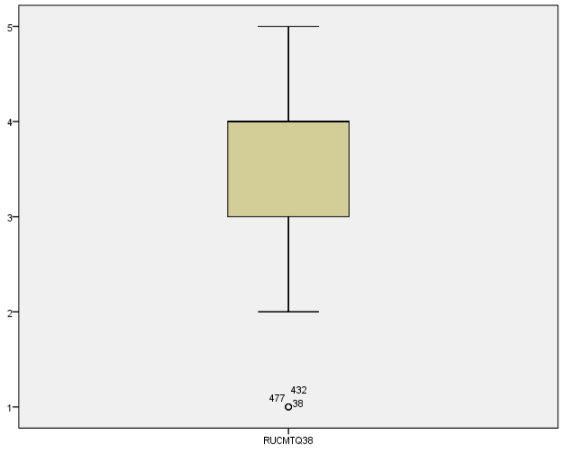

481 436 414
313
RUCMTQ33

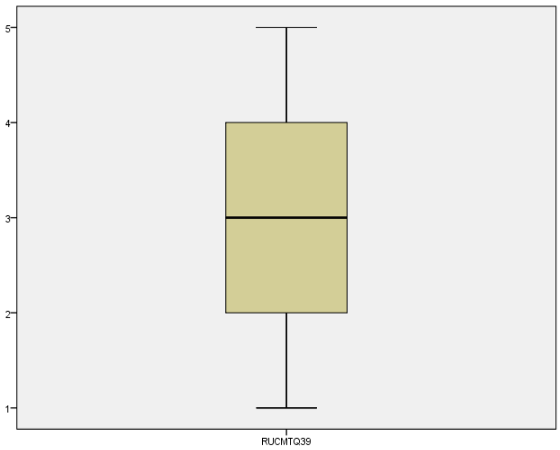

459 473
465
464
457 469
462
418
467 432
414
UCMTQ34

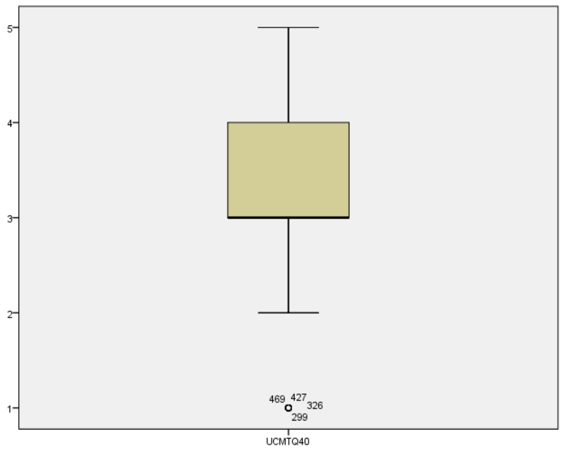

448 366 302
147
RUCMTQ35

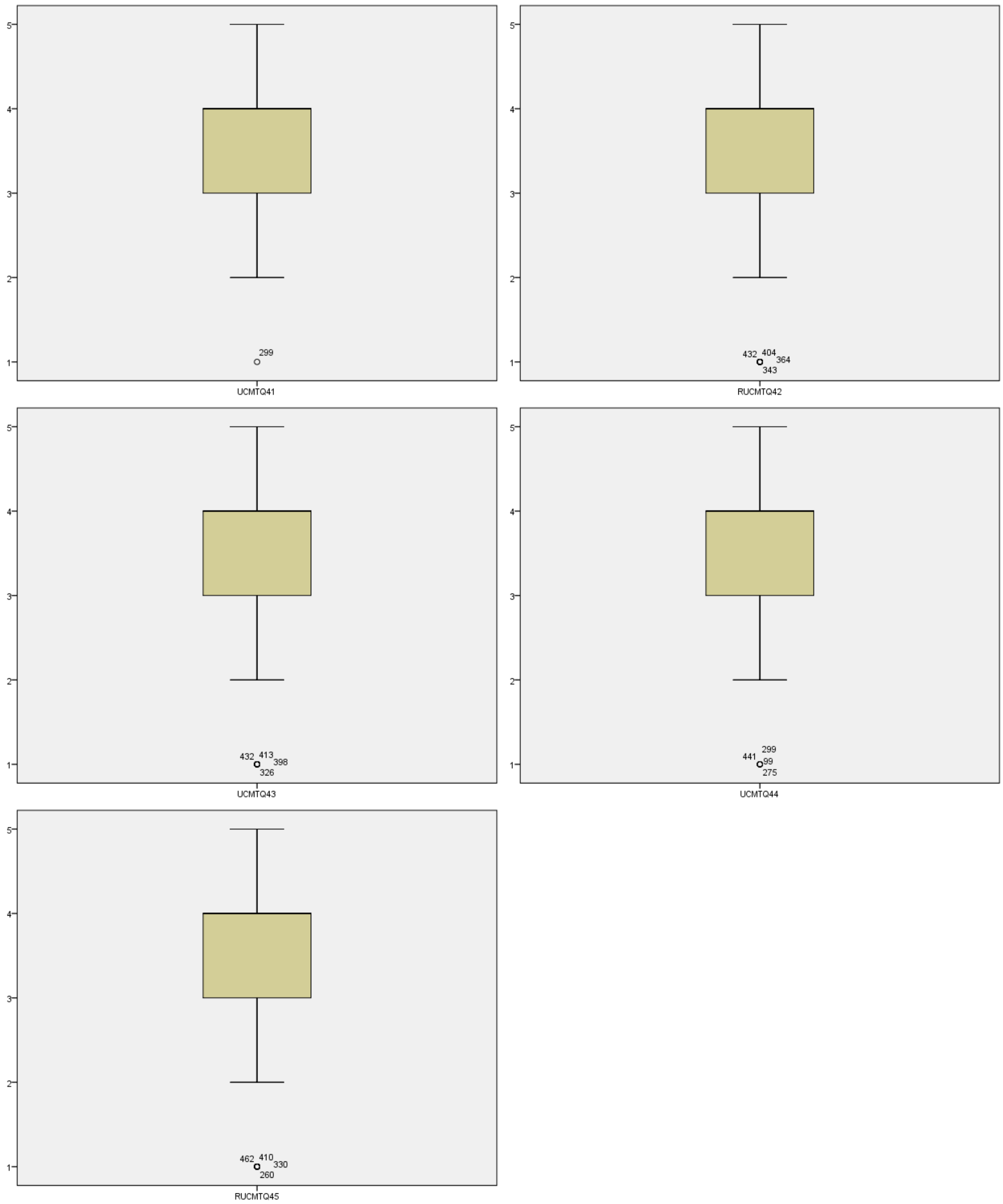

450 441 432
404
RUCMTQ36


460 479
478
475
433 445
435
334
360 357
270
UCMTQ37


432
477 38
RUCMTQ38


RUCMTQ39


469 427 326
299
UCMTQ40


UCMTQ41
RUCMTQ42
UCMTQ43
UCMTQ44
RUCMTQ45

Appendix 2.5
Multivariate outliers (Mahalanobis distance)

Observation number	Mahalanobis d-squared	p1	p2
300	139.697	.000	.000
80	116.151	.000	.000
318	116.017	.000	.000
357	115.670	.000	.000
121	111.662	.000	.000
441	97.870	.000	.000
275	93.581	.000	.000
82	92.386	.000	.000
32	89.906	.000	.000
109	89.323	.000	.000
227	87.747	.000	.000
247	87.085	.000	.000
214	86.372	.000	.000
58	85.942	.000	.000
448	84.519	.000	.000
268	83.987	.000	.000
218	82.665	.001	.000
260	82.062	.001	.000
125	81.982	.001	.000
299	81.046	.001	.000
52	80.775	.001	.000
197	80.417	.001	.000
168	79.333	.001	.000
100	78.859	.001	.000
356	78.739	.001	.000
164	78.109	.002	.000
233	78.092	.002	.000
181	77.994	.002	.000
281	77.929	.002	.000
237	77.234	.002	.000
170	77.054	.002	.000
430	76.827	.002	.000
238	75.967	.003	.000
271	74.618	.004	.000
189	74.107	.004	.000
8	73.821	.004	.000
103	73.562	.005	.000
239	73.471	.005	.000
386	72.694	.006	.000
47	72.469	.006	.000
270	72.426	.006	.000
49	72.055	.006	.000
366	71.850	.007	.000

Observation number	Mahalanobis d-squared	p1	p2
187	70.907	.008	.000
343	70.851	.008	.000
411	69.740	.010	.000
241	69.554	.011	.000
403	69.118	.012	.000
145	69.009	.012	.000
477	68.896	.012	.000
413	68.861	.013	.000
405	68.661	.013	.000
213	68.229	.014	.000
360	67.471	.017	.000
205	67.370	.017	.000
370	66.950	.018	.000
396	66.794	.019	.000
340	66.380	.021	.000
330	66.145	.022	.000
432	65.791	.023	.000
61	65.429	.025	.000
462	65.082	.027	.000
328	64.928	.027	.000
202	64.794	.028	.000
217	64.778	.028	.000
23	64.712	.029	.000
98	64.451	.030	.000
138	64.328	.031	.000
211	64.226	.031	.000
128	64.189	.032	.000
474	63.736	.034	.000
230	63.575	.035	.000
9	63.462	.036	.000
465	63.139	.038	.000
246	62.801	.041	.000
33	62.664	.042	.000
60	62.512	.043	.000
182	62.100	.046	.000
334	62.073	.046	.000
22	61.720	.049	.000
304	61.715	.049	.000
415	61.604	.050	.000
195	61.075	.055	.000
206	61.070	.055	.000
89	61.057	.056	.000
293	60.771	.058	.000
51	60.658	.060	.000
463	60.653	.060	.000
451	60.348	.063	.000

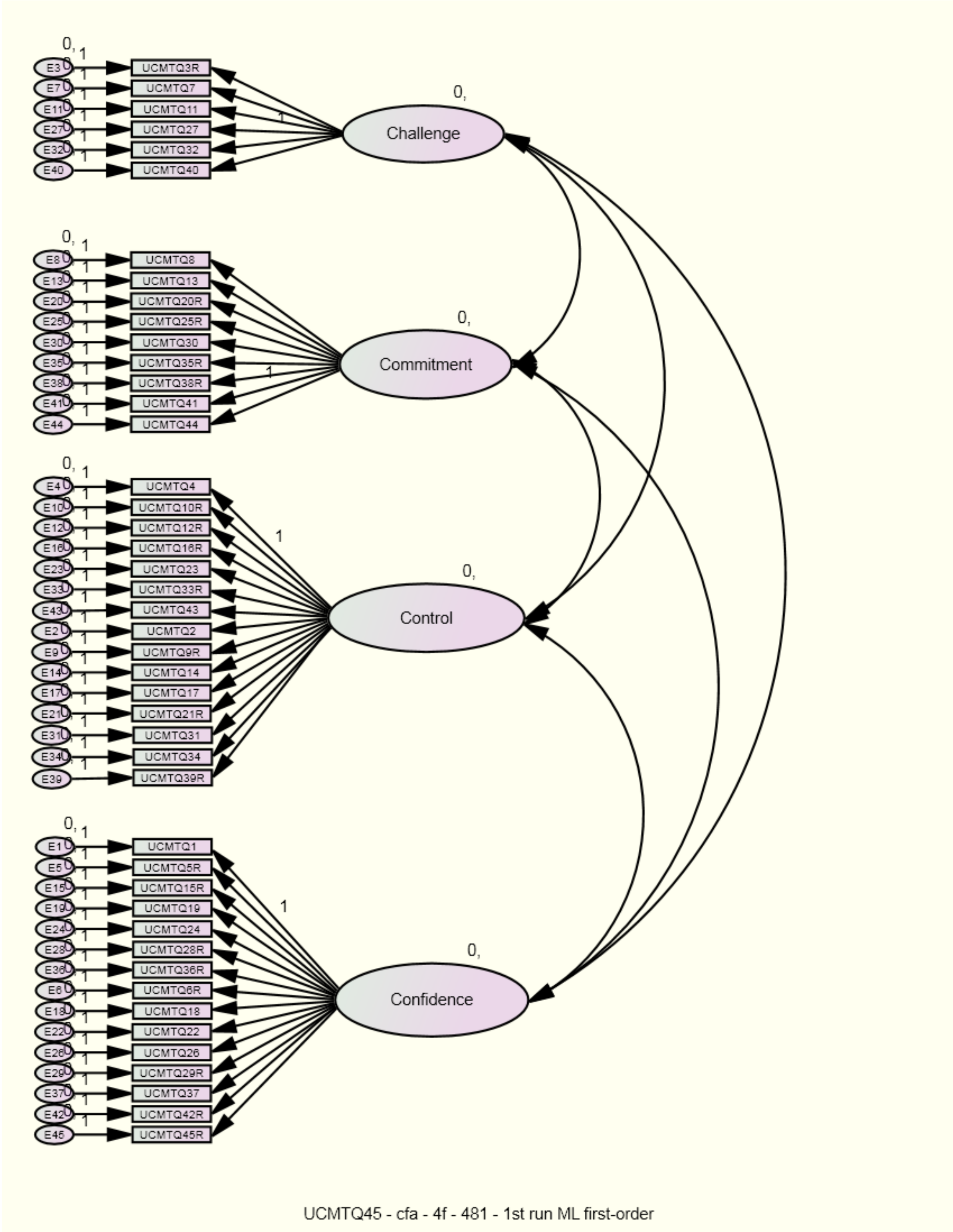
Observation number	Mahalanobis d-squared	p1	p2
265	60.231	.064	.000
433	60.163	.065	.000
249	60.161	.065	.000
453	60.089	.066	.000
29	60.071	.066	.000
78	59.997	.067	.000
425	59.807	.069	.000
120	59.758	.069	.000
236	59.569	.072	.000
18	59.131	.077	.000
71	58.711	.082	.000

Appendix 2.6
Univariate normality statistics - extreme outliers removed (data set B)

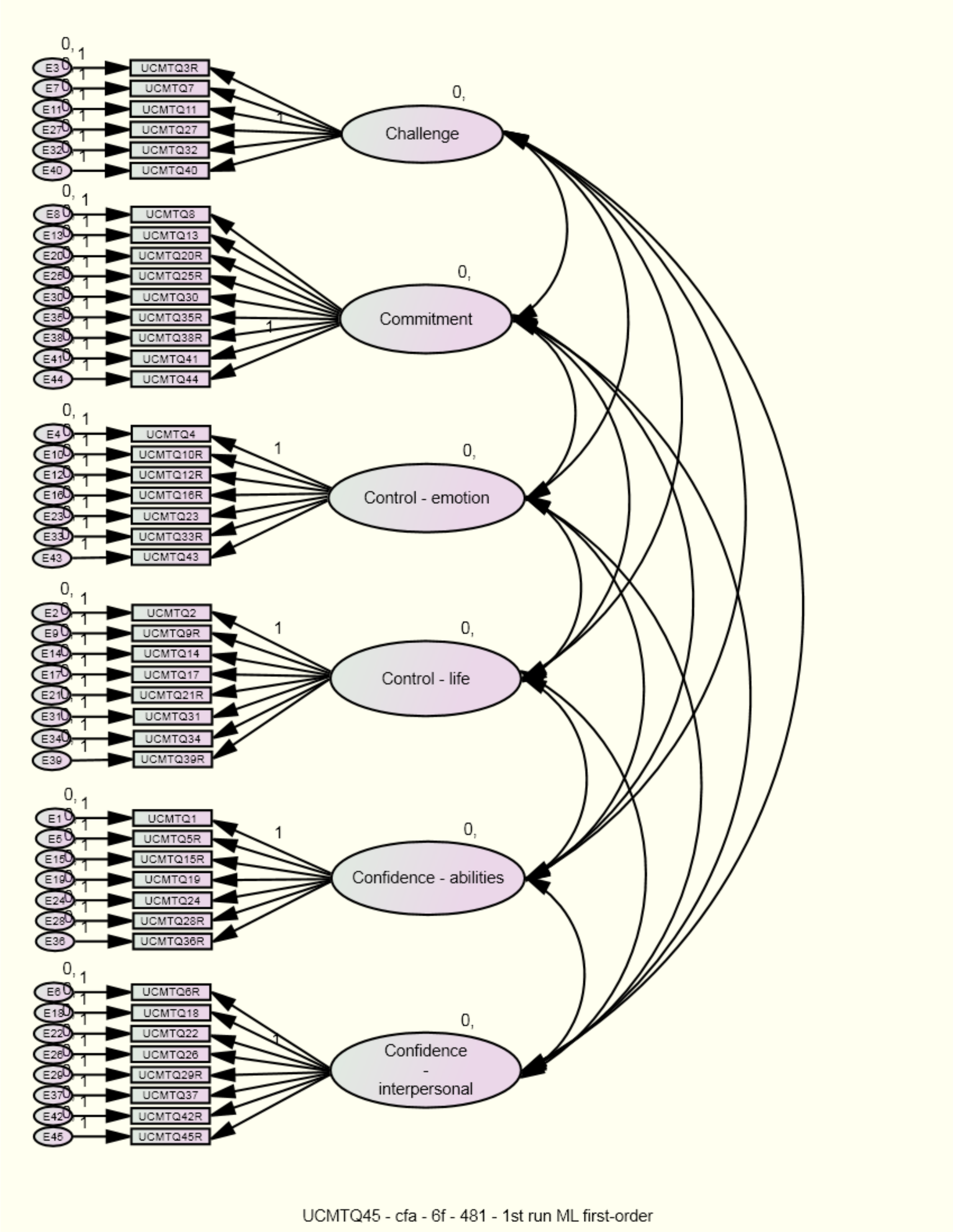
	Skewness ratio	Kurtosis ratio
UCMTQ1	-7.78	4.47
UCMTQ2	-5.31	4.68
RUCMTQ3	-3.91	-2.09
UCMTQ4	-3.41	-2.51
RUCMTQ5	0.40	-4.00
RUCMTQ6	-3.23	-1.57
UCMTQ7	2.03	-2.44
UCMTQ8	-6.45	4.33
RUCMTQ9	3.60	-3.35
RUCMTQ10	-2.44	-3.83
UCMTQ11	-6.68	5.36
RUCMTQ12	-3.35	-2.05
UCMTQ13	-6.13	3.31
UCMTQ14	-4.66	1.34
RUCMTQ15	-1.51	-4.14
RUCMTQ16	-2.49	-3.36
UCMTQ17	-5.28	7.39
UCMTQ18	-6.68	1.68
UCMTQ19	-8.11	5.84
RUCMTQ20	-2.93	-3.79
RUCMTQ21	4.00	-2.07
UCMTQ22	-5.42	1.05
UCMTQ23	-6.09	-0.36
UCMTQ24	-6.82	5.77
RUCMTQ25	-5.59	2.07
UCMTQ26	-6.65	6.55
UCMTQ27	-4.58	1.27
RUCMTQ28	3.73	-2.91
RUCMTQ29	-6.18	2.11
UCMTQ30	-4.45	3.88
UCMTQ31	-1.96	-2.57
UCMTQ32	-4.00	-0.57
RUCMTQ33	4.69	-2.07
UCMTQ34	-3.64	2.30
RUCMTQ35	-5.86	-0.04
RUCMTQ36	-4.60	-2.16
UCMTQ37	-1.74	2.38

RUCMTQ38	-5.79	1.17
RUCMTQ39	-2.06	-3.91
UCMTQ40	-2.36	0.10
UCMTQ41	-5.20	2.34
RUCMTQ42	-3.22	-2.70
UCMTQ43	-2.15	-3.52
UCMTQ44	-5.09	-0.12
RUCMTQ45	-4.32	-2.05

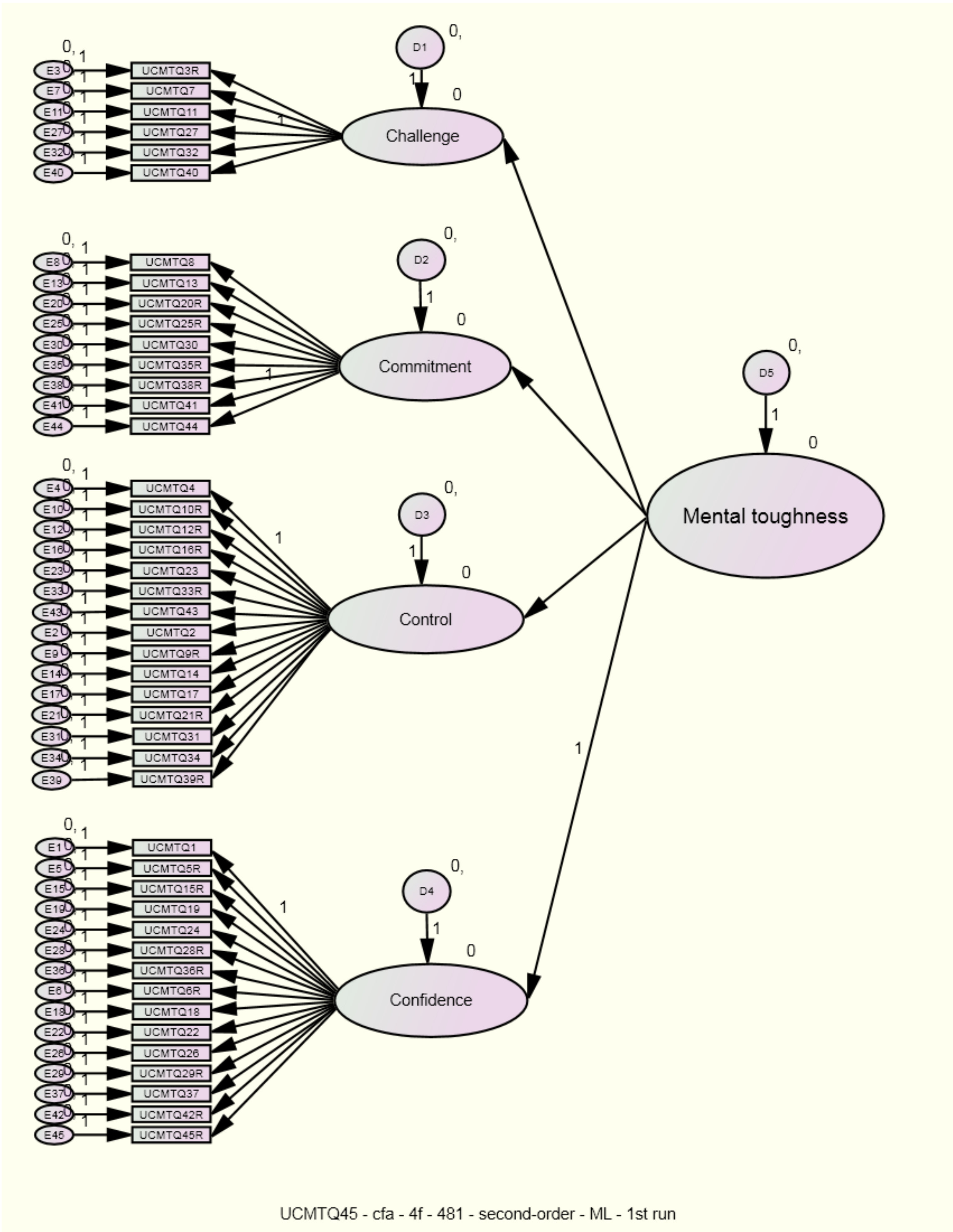
Appendix 2.7
Path diagrams of the respective specified models of the UCMTQ



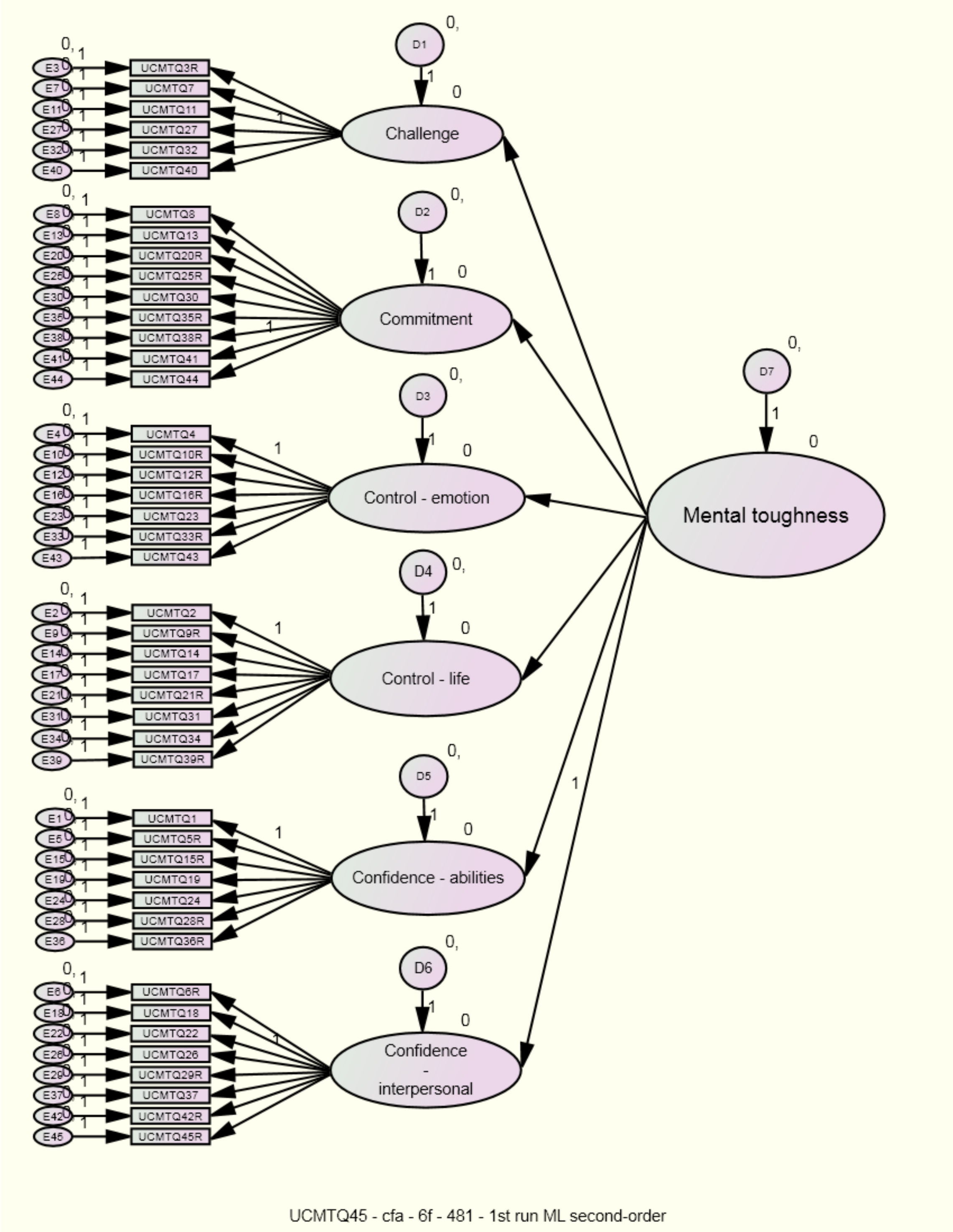
AMOS path diagram depicting first-order four factor model of mental toughness



AMOS path diagram depicting first-order six factor model of mental toughness



AMOS path diagram depicting second-order four factor model of mental toughness



AMOS path diagram depicting second-order six factor model of mental toughness

Appendix 2.8
Regression weights of the respective models of the UCMTQ

Standardised regression weights of the first-order four factor model

Factor / item			Estimate
UCMTQ27	<---	Challenge	0.725*
UCMTQ32	<---	Challenge	0.590*
UCMTQ11	<---	Challenge	0.533*
UCMTQ40	<---	Challenge	0.456*
UCMTQ3R	<---	Challenge	0.396*
UCMTQ7	<---	Challenge	0.235*
UCMTQ30	<---	Commitment	0.685*
UCMTQ38R	<---	Commitment	0.676*
UCMTQ25R	<---	Commitment	0.662*
UCMTQ13	<---	Commitment	0.662*
UCMTQ41	<---	Commitment	0.651*
UCMTQ8	<---	Commitment	0.449*
UCMTQ20R	<---	Commitment	0.399*
UCMTQ35R	<---	Commitment	0.330*
UCMTQ44	<---	Commitment	0.319*
UCMTQ12R	<---	Control	0.759*
UCMTQ4	<---	Control	0.646*
UCMTQ16R	<---	Control	0.616*
UCMTQ10R	<---	Control	0.611*
UCMTQ43	<---	Control	0.550*
UCMTQ2	<---	Control	0.530*
UCMTQ39R	<---	Control	0.499*
UCMTQ17	<---	Control	0.393*
UCMTQ14	<---	Control	0.329*
UCMTQ33R	<---	Control	0.280*
UCMTQ34	<---	Control	0.217*
UCMTQ21R	<---	Control	0.093
UCMTQ31	<---	Control	0.090
UCMTQ9R	<---	Control	0.040
UCMTQ23	<---	Control	-0.057
UCMTQ36R	<---	Confidence	0.741*
UCMTQ15R	<---	Confidence	0.699*
UCMTQ1	<---	Confidence	0.655*

UCMTQ5R	<---	Confidence	0.638*
UCMTQ19	<---	Confidence	0.575*
UCMTQ42R	<---	Confidence	0.443*
UCMTQ28R	<---	Confidence	0.442*
UCMTQ24	<---	Confidence	0.442*
UCMTQ6R	<---	Confidence	0.427*
UCMTQ29R	<---	Confidence	0.411*
UCMTQ18	<---	Confidence	0.387*
UCMTQ26	<---	Confidence	0.366*
UCMTQ22	<---	Confidence	0.342*
UCMTQ37	<---	Confidence	0.252*
UCMTQ45R	<---	Confidence	0.229*

Note: * Indicates significant relationship ($p < 0.05$).

Standardised regression weights of the first-order six factor model

Factor / item			Estimate
UCMTQ27	<---	Challenge	0.727*
UCMTQ32	<---	Challenge	0.589*
UCMTQ11	<---	Challenge	0.535*
UCMTQ40	<---	Challenge	0.454*
UCMTQ3R	<---	Challenge	0.395*
UCMTQ7	<---	Challenge	0.236*
UCMTQ30	<---	Commitment	0.690*
UCMTQ38R	<---	Commitment	0.675*
UCMTQ13	<---	Commitment	0.662*
UCMTQ25R	<---	Commitment	0.655*
UCMTQ41	<---	Commitment	0.650*
UCMTQ8	<---	Commitment	0.452*
UCMTQ20R	<---	Commitment	0.398*
UCMTQ35R	<---	Commitment	0.329*
UCMTQ44	<---	Commitment	0.321*
UCMTQ12R	<---	Control - emotion	0.809*
UCMTQ4	<---	Control - emotion	0.673*
UCMTQ16R	<---	Control - emotion	0.639*
UCMTQ10R	<---	Control - emotion	0.632*
UCMTQ43	<---	Control - emotion	0.563*
UCMTQ33R	<---	Control - emotion	0.290*
UCMTQ23	<---	Control - emotion	-0.043
UCMTQ2	<---	Control - life	0.630*
UCMTQ17	<---	Control - life	0.553*
UCMTQ14	<---	Control - life	0.496*
UCMTQ39R	<---	Control - life	0.466*
UCMTQ34	<---	Control - life	0.358*
UCMTQ31	<---	Control - life	0.203*
UCMTQ21R	<---	Control - life	0.045
UCMTQ9R	<---	Control - life	0.006
UCMTQ36R	<---	Confidence - abilities	0.752*
UCMTQ15R	<---	Confidence - abilities	0.722*
UCMTQ5R	<---	Confidence - abilities	0.670*

UCMTQ1	<---	Confidence - abilities	0.660*
UCMTQ19	<---	Confidence - abilities	0.582*
UCMTQ28R	<---	Confidence - abilities	0.443*
UCMTQ24	<---	Confidence - abilities	0.432*
UCMTQ42R	<---	Confidence - interpersonal	0.635*
UCMTQ18	<---	Confidence - interpersonal	0.586*
UCMTQ6R	<---	Confidence - interpersonal	0.481*
UCMTQ26	<---	Confidence - interpersonal	0.450*
UCMTQ29R	<---	Confidence - interpersonal	0.442*
UCMTQ37	<---	Confidence - interpersonal	0.394*
UCMTQ22	<---	Confidence - interpersonal	0.393*
UCMTQ45R	<---	Confidence - interpersonal	0.321*

Note: * Indicates significant relationship ($p < 0.05$).

Standardised regression weights of the second-order four factor model

Factor / item			Estimate
Confidence	<---	Mental toughness	0.941*
Control	<---	Mental toughness	0.861*
Commitment	<---	Mental toughness	0.763*
Challenge	<---	Mental toughness	0.746*
UCMTQ27	<---	Challenge	0.683*
UCMTQ32	<---	Challenge	0.641*
UCMTQ11	<---	Challenge	0.537*
UCMTQ40	<---	Challenge	0.496*
UCMTQ3R	<---	Challenge	0.352*
UCMTQ7	<---	Challenge	0.227*
UCMTQ38R	<---	Commitment	0.705*
UCMTQ25R	<---	Commitment	0.683*
UCMTQ41	<---	Commitment	0.657*
UCMTQ30	<---	Commitment	0.652*
UCMTQ13	<---	Commitment	0.639*
UCMTQ8	<---	Commitment	0.418*
UCMTQ20R	<---	Commitment	0.409*
UCMTQ35R	<---	Commitment	0.355*
UCMTQ44	<---	Commitment	0.319*
UCMTQ12R	<---	Control	0.751*
UCMTQ4	<---	Control	0.650*
UCMTQ10R	<---	Control	0.612*
UCMTQ16R	<---	Control	0.608*
UCMTQ43	<---	Control	0.550*
UCMTQ2	<---	Control	0.530*
UCMTQ39R	<---	Control	0.500*
UCMTQ17	<---	Control	0.403*
UCMTQ14	<---	Control	0.342*
UCMTQ33R	<---	Control	0.274*
UCMTQ34	<---	Control	0.230*
UCMTQ31	<---	Control	0.096
UCMTQ21R	<---	Control	0.090
UCMTQ9R	<---	Control	0.037
UCMTQ23	<---	Control	-0.048

UCMTQ36R	<---	Confidence	0.736*
UCMTQ15R	<---	Confidence	0.695*
UCMTQ1	<---	Confidence	0.656*
UCMTQ5R	<---	Confidence	0.628*
UCMTQ19	<---	Confidence	0.580*
UCMTQ24	<---	Confidence	0.450*
UCMTQ42R	<---	Confidence	0.448*
UCMTQ28R	<---	Confidence	0.438*
UCMTQ6R	<---	Confidence	0.435*
UCMTQ29R	<---	Confidence	0.409*
UCMTQ18	<---	Confidence	0.395*
UCMTQ26	<---	Confidence	0.364*
UCMTQ22	<---	Confidence	0.350*
UCMTQ37	<---	Confidence	0.259*
UCMTQ45R	<---	Confidence	0.233*

Note: * Indicates significant relationship ($p < 0.05$).

Standardised regression weights of the second-order six factor model

	Factor / item		Estimate
Confidence - abilities	<---	Mental toughness	0.889*
Control - life	<---	Mental toughness	0.836*
Commitment	<---	Mental toughness	0.794*
Challenge	<---	Mental toughness	0.781*
Control - emotion	<---	Mental toughness	0.747*
Confidence - interpersonal	<---	Mental toughness	0.704*
UCMTQ27	<---	Challenge	0.686*
UCMTQ32	<---	Challenge	0.637*
UCMTQ11	<---	Challenge	0.539*
UCMTQ40	<---	Challenge	0.494*
UCMTQ3R	<---	Challenge	0.356*
UCMTQ7	<---	Challenge	0.224*
UCMTQ38R	<---	Commitment	0.703*
UCMTQ25R	<---	Commitment	0.678*
UCMTQ30	<---	Commitment	0.657*
UCMTQ41	<---	Commitment	0.656*
UCMTQ13	<---	Commitment	0.641*
UCMTQ8	<---	Commitment	0.420*
UCMTQ20R	<---	Commitment	0.408*
UCMTQ35R	<---	Commitment	0.353*
UCMTQ44	<---	Commitment	0.320*
UCMTQ12R	<---	Control - emotion	0.804*
UCMTQ4	<---	Control - emotion	0.680*
UCMTQ10R	<---	Control - emotion	0.640*
UCMTQ16R	<---	Control - emotion	0.635*
UCMTQ43	<---	Control - emotion	0.562*
UCMTQ33R	<---	Control - emotion	0.283*
UCMTQ23	<---	Control - emotion	-0.034
UCMTQ2	<---	Control - life	0.638*
UCMTQ17	<---	Control - life	0.538*

UCMTQ14	<---	Control - life	0.507*
UCMTQ39R	<---	Control - life	0.458*
UCMTQ34	<---	Control - life	0.365*
UCMTQ31	<---	Control - life	0.201*
UCMTQ21R	<---	Control - life	0.048
UCMTQ9R	<---	Control - life	0.008
UCMTQ36R	<---	Confidence - abilities	0.752*
UCMTQ15R	<---	Confidence - abilities	0.719*
UCMTQ1	<---	Confidence - abilities	0.669*
UCMTQ5R	<---	Confidence - abilities	0.656*
UCMTQ19	<---	Confidence - abilities	0.595*
UCMTQ24	<---	Confidence - abilities	0.439*
UCMTQ28R	<---	Confidence - abilities	0.431*
UCMTQ42R	<---	Confidence - interpersonal	0.647*
UCMTQ18	<---	Confidence - interpersonal	0.600*
UCMTQ6R	<---	Confidence - interpersonal	0.478*
UCMTQ26	<---	Confidence - interpersonal	0.443*
UCMTQ29R	<---	Confidence - interpersonal	0.420*
UCMTQ22	<---	Confidence - interpersonal	0.401*
UCMTQ37	<---	Confidence - interpersonal	0.388*
UCMTQ45R	<---	Confidence - interpersonal	0.325*

Note: * Indicates significant relationship ($p < 0.05$).

Appendix 2.9
Independent model re-specifications of the respective models of the University of Chichester Mental Toughness Questionnaire

UCMTQ fit indices of the first-order four factor model re-specification

		Fit indices					
CFA run	Original model	CMIN/DF	*P*	CFI	PCFI	RMSEA (90% CI)	AIC
Criterion values	removed	(< 2.00)	(> 0.05)	(> 0.95)	(> 0.6)	(< 0.05)	(lower = better)
1	Original model	3.115	0.000	0.681	0.646*	0.066 (0.064, 0.069)	3206.887
2	21R, 9R	2.690	0.000	0.746	0.705*	0.059 (0.056, 0.062)	2567.476
3	18	2.588	0.000	0.764	0.721*	0.058 (0.055, 0.061)	2368.194
4	14, 34	2.533	0.000	0.785	0.739*	0.057 (0.053, 0.060)	2111.195
5	37, 26	2.514	0.000	0.802	0.751*	0.056 (0.053, 0.060)	1896.976
6	17	2.454	0.000	0.815	0.762*	0.055 (0.052, 0.059)	1762.625
7	22	2.396	0.000	0.827	0.772*	0.054 (0.050, 0.058)	1637.054
8	28R	2.401	0.000	0.832	0.774*	0.054 (0.050, 0.058)	1552.037
9	3R, 7, 35R, 44, 33R, 31, 23, 45R, 29R	2.564	0.000	0.885	0.798*	0.057 (0.052, 0.062)	919.234

* indicates good fit.

UCMTQ fit indices of the first-order six factor model re-specification

		Fit indices					
CFA run	Items removed	CMIN/DF	*P*	CFI	PCFI	RMSEA	AIC
Criterion values		(< 2.00)	(> 0.05)	(> 0.95)	(> 0.6)	(< 0.05)	(lower = better)
1	Original model	2.830	0.000	0.727	0.683*	0.062 (0.059, 0.065)	2931.694
2	21R 9R	2.368	0.000	0.796	0.745*	0.053 (0.050, 0.056)	2289.056
3	26	2.333	0.000	0.806	0.752*	0.053 (0.050, 0.056)	2157.765
4	34	2.367	0.000	0.808	0.753*	0.053 (0.050, 0.057)	2084.546
5	39R	2.342	0.000	0.816	0.758*	0.053 (0.050, 0.056)	1968.125
6	28R	2.341	0.000	0.821	0.761*	0.053 (0.050, 0.056)	1872.332
7	3R, 7, 20R, R35, 44, 33R, 23, 31, 37, 45R	2.738	0.000	0.860	0.767*	0.060 (0.056, 0.065)	1195.333

* indicates good fit.

UCMTQ fit indices of the second-order four factor model re-specification

		Fit indices					
CFA run	Items removed	CMIN/DF	*P*	CFI	PCFI	RMSEA	AIC
Criterion values		(< 2.00)	(> 0.05)	(> 0.95)	(> 0.6)	(< 0.05)	(lower = better)
1	Original model	3.238	0.000	0.622	0.629*	0.068 (0.066, 0.071)	3325.418
2	21R, 9R,	2.827	0.000	0.725	0.687*	0.062 (0.059, 0.065)	2685.635
3	18	2.735	0.000	0.741	0.701*	0.060 (0.057, 0.063)	2488.968
4	14, 34	2.703	0.000	0.761	0.718*	0.060 (0.056, 0.063)	2237.060
5	37, 26	2.703	0.000	0.776	0.730*	0.060 (0.056, 0.063)	2022.780
6	22	2.671	0.000	0.787	0.738*	0.059 (0.056, 0.062)	1899.316
7	28R	2.691	0.000	0.790	0.740*	0.059 (0.056, 0.063)	1811.634
8	3R, 7, 35, 44, 33R, 31, 23, 45R, 29R, 17	2.942	0.000	0.857	0.778*	0.064 (0.059, 0.069)	1031.807

* indicates good fit.

UCMTQ fit indices of the second-order six factor model re-specification

		Fit indices					
CFA run	Items removed	CMIN/DF	*P*	CFI	PCFI	RMSEA	AIC
Criterion values		(< 2.00)	(> 0.05)	(> 0.95)	(> 0.6)	(< 0.05)	(lower = better)
1	Original model	2.961	0.000	0.704	0.668*	0.064 (0.061, 0.067)	3062.289
2	21R, 9R	2.517	0.000	0.772	0.730*	0.056 (0.053, 0.059)	2419.757
3	37	2.537	0.000	0.776	0.733*	0.057 (0.054, 0.060)	2326.847
4	34	2.583	0.000	0.778	0.733*	0.057 (0.054, 0.061)	2254.486
5	39R	2.556	0.000	0.785	0.739*	0.057 (0.054, 0.060)	2135.408
6	28R	2.568	0.000	0.790	0.742*	0.057 (0.054, 0.060)	2033.154
7	3R, 7, 35R, 44, 33R, 23, 31, 45R	2.917	0.000	0.825	0.760*	0.063 (0.059, 0.067)	1446.340

* indicates good fit.

Study 3 Assessment Instruments and SPSS/AMOS Output

Appendix 3.1
The Hardiness Confidence Questionnaire

Considering how you are **generally**, please answer the following statements. Start each statement with the phrase “In my general life.” For example, the first statement should be read “In my general life, I am confident in my own abilities.”

Please answer **honestly**, we are looking for your initial response to each statement so try not to spend too much time on any one statement.

Please indicate your response by **circling one** of the numbers, which have the following meaning;

1 = strongly disagree; **2** = disagree; **3** = neither agree nor disagree; **4** = agree; **5** = strongly agree.

In my general life, …	Strongly disagree	Disagree	Neither agree or disagree	Agree	Strongly agree
1. I am confident in my own abilities	1	2	3	4	5
2. I am eager to stay dedicated to a task	1	2	3	4	5
3. I do not look forward to challenging situations	1	2	3	4	5
4. I can influence the path that my life takes	1	2	3	4	5
5. I doubt whether I am a worthwhile person	1	2	3	4	5
6. I often find myself disengaging from tasks	1	2	3	4	5
7. I expect to succeed in performing important tasks	1	2	3	4	5
8. I find it difficult to stay committed to whatever I am doing	1	2	3	4	5
9. I doubt myself when I have a difficult task to undertake	1	2	3	4	5
10. I believe that change enables me to grow as a person	1	2	3	4	5
11. I find it difficult to stay committed to achieving my goals	1	2	3	4	5
12. I am an influential person	1	2	3	4	5
13. I see challenges in my life as opportunities for me to develop as a person	1	2	3	4	5
14. Events in my life are determined by others	1	2	3	4	5
15. I lack self-belief	1	2	3	4	5
16. Events in my life are shaped by my own actions	1	2	3	4	5
17. I feel that I am a person of worth	1	2	3	4	5
18. I stay committed to tasks even in the face of difficulty	1	2	3	4	5

In my general life, …	Strongly disagree	Disagree	Neither agree or disagree	Agree	Strongly agree
19. My actions have little influence on my life	1	2	3	4	5
20. Changes in my daily routine encourage me to learn	1	2	3	4	5
21. I take a negative attitude toward myself	1	2	3	4	5
22. What happens to me is my own doing	1	2	3	4	5
23. I lack commitment	1	2	3	4	5
24. I do not like to make changes to my everyday schedule	1	2	3	4	5
25. I feel that I am a truly worthwhile person	1	2	3	4	5
26. I believe that change is a normal part of life	1	2	3	4	5
27. I feel that I have little control over the direction my life is taking	1	2	3	4	5

Appendix 3.2
The Hardiness Confidence Questionnaire Scoring Key

Constructs	Items	Total number of items
Challenge	3R, 10, 13, 20, 24R, 26	6 (2 X R)
Commitment	2, 6R, 8R, 11R, 18, 23R,	6 (4 X R)
Control	4, 12, 14R, 16, 19R, 22, 27R	7 (3 X R)
Confidence - competence	1, 7, 9R, 15R	4 (2 X R)
Confidence – self worth	5R, 17, 21R, 25	4 (2 X R)

Calculating scores

Total mean MT = Add all 27 items scores (accounting for 13 reversed items) and divide by 27.

Mean construct scores = Add the respective items scores (accounting for reversed items!) for each construct and divide by the number of items for that construct.

Appendix 3.3
Participant information sheet, consent form, demographic questionnaire and written debrief

Department of Sport and Exercise Sciences
Project information sheet for participants

PLEASE READ THE FOLLOWING CAREFULLY
Study title: An examination of the psychological characteristics and thought processes of competitive athletes.

What is the purpose of the research and how will the research be carried out?
The purpose of this study is to examine the psychological characteristics and thought processes of competitive athletes.

At the University of Chichester before any project starts it has to be checked by university staff. They make sure that the research is good enough to carry out. This project has been checked by the staff of the Department of Sport and Exercise Science and has been passed by the University Ethics Committee.

What will you be asked to do?
The study involves completing two questionnaires which will assess your psychological characteristics and thought processes. Please take your time to complete the questionnaires and answer as truthfully as possible.

In total, each questionnaire should take no longer than 5 minutes so the total amount of time you will need to commit to this project is 10 minutes.

What are the anticipated benefits of participating in the research?
You have been asked to take part in this project since we need athletes who are competing in sport to gain an insight into the psychology of sport. It is hoped that the research will enhance our knowledge of what attributes contribute to sporting success.

Are there any risks associated with participating in the research?
You will be providing potentially sensitive information about perceptions of your psychological characteristics and thought processes. As such, while the information you provide will be used in academic publications, your name will not be mentioned in any literature, and as such, the responses you make would not be linked back to you.

Do you have to take part?
No! It is your choice whether you participate in the project or not. If at any time during the project you feel that you don't want to continue then you can tell the researcher that you want to stop – you do not have to give a reason.

What will happen to the information collected as part of the study?
All the information which is collected about you during the project will be kept private and stored in lockable cabinets – no one will know that the information belongs to you and all the information will be kept at the University of Chichester. Your responses and information will be treated confidentially. Questionnaire responses will be associated with an anonymous code rather than a person's name. The questionnaires will be stored for a period of 7 years and then destroyed. If you wish to withdraw your data from the study you should do so within 6 months of completing the study.

Who can you contact if you have a complaint about the project?
Dr Andy Dixon – Head of the Research and Employer Engagement Office – Email a.dixon@chi.ac.uk: Phone 01243 812125

Who can you contact if you have any questions about the project?
Phil Birch – Email p.birch@chi.ac.uk: Phone 01243 816345 or
Dr Iain Greenlees - Email i.greenlees@chi.ac.uk: Phone 01243 816437
Version 1, 15/04/2012 **Thank you for your time**

Participant consent form

PLEASE READ THE FOLLOWING CAREFULLY

Study title: An examination of the psychological characteristics and thought processes of competitive athletes.

I understand that my participation in this project will involve completing 2 questionnaires which will assess my perceptions of psychological characteristics and thought processes.

I am happy to participate in this research Yes No

1) I have read and understand the information sheet for this research project. I have had the opportunity to consider the information, ask questions and have had these answered satisfactorily. Yes No

2) I understand that my participation in the activity is voluntary and that I am therefore free to withdraw my involvement at any stage, without giving a reason. Yes No

3) I understand that all information will be anonymised and that my personal information will not be released to any third parties. Yes No

If you responded "Yes" to the above statements please complete the following:

Your name (please print)…………………………………………………

Your signature…………………………………………………………………

Date..

Thank you for your time

Version 1, 15/04/12

Demographic Questionnaire

Please fill in your details in the spaces provided

Age ______________ Full initials _______________ Date of Birth ________ / ________ / ________

(years)

Gender Male ☐ Female ☐

Ethnic origin – Please indicate the most appropriate description of your ethnic origin.

(Please tick)

White - British	☐	Asian - Indian	☐
White - Irish	☐	Asian - Pakistani	☐
White – Any other White background	☐	Asian – Bangladeshi	☐
Mixed – White & Black Caribbean	☐	Black or Black British - Caribbean	☐
Mixed – White & Black African	☐	Black or Black British - African	☐
Mixed – White & Black Asian	☐	Black or Black British - Any other Black background	☐
Mixed – Any other Mixed background	☐	Chinese – Any Chinese background	☐
Any other ethnic background	☐		

University of study ____

Year of study ____

(e.g., 1, 2, 3)

What is your primary sport? ____

How many years competitive playing experience do you have in your primary sport? ___________________________

What is the highest level that you have played in your primary sport? _____________________________________

Participant debrief

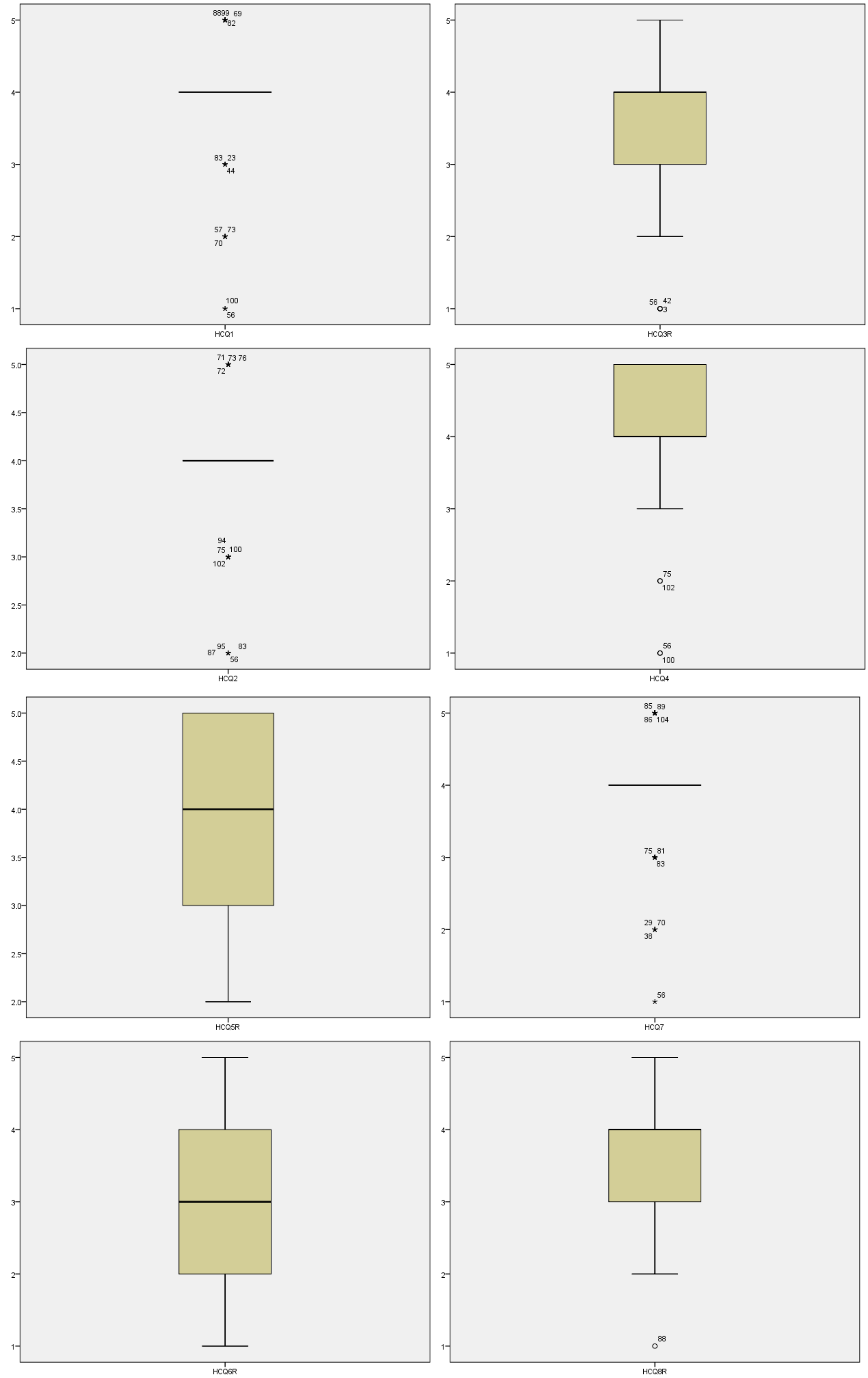
The aim of this study was to examine the suitability of Hardiness Confidence Questionnaire in measuring mental toughness in sport and to devise programmes to enhance the suitability of the HCQ.

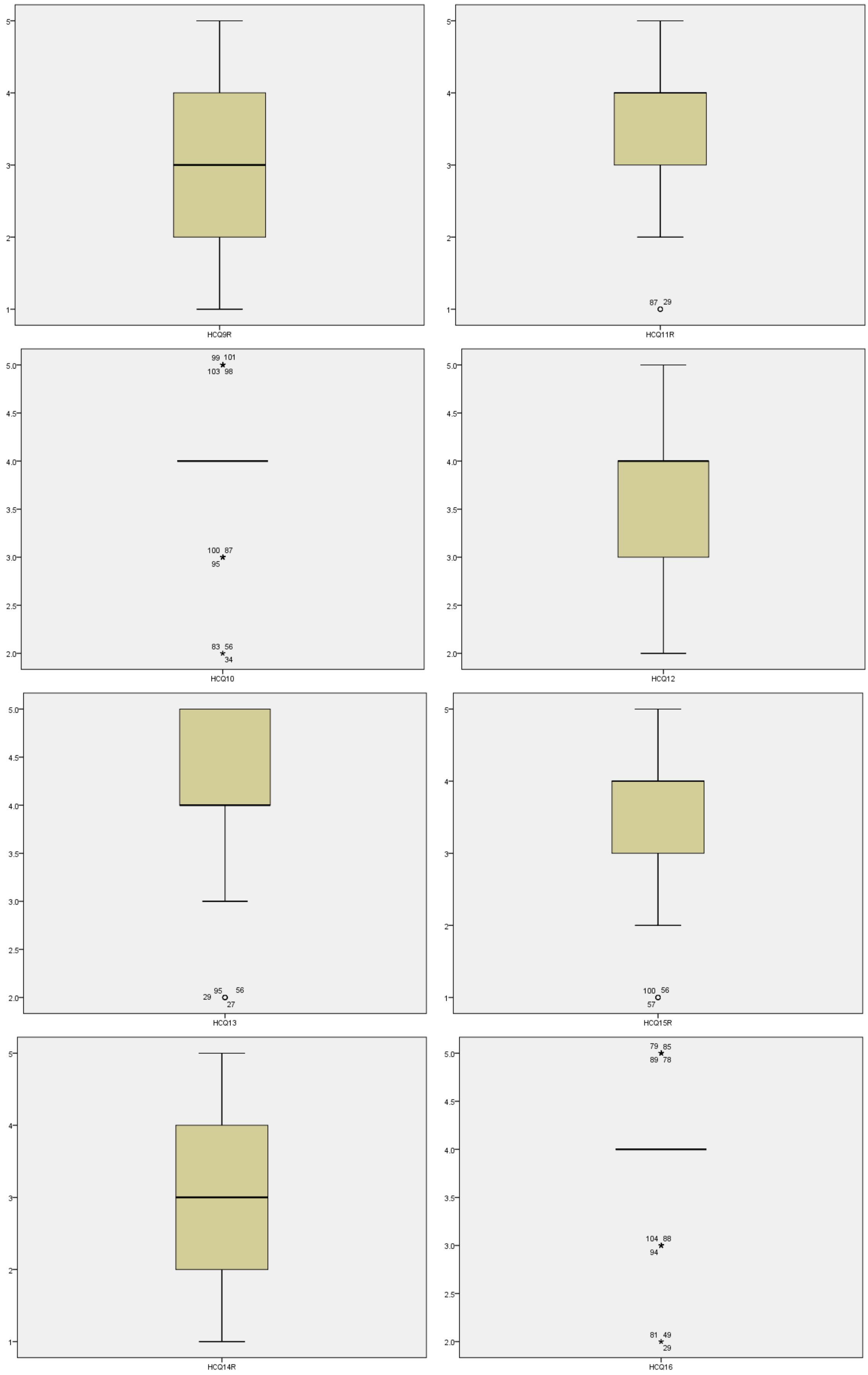
This required us to compare the HCQ with a Social Desirability Scale.

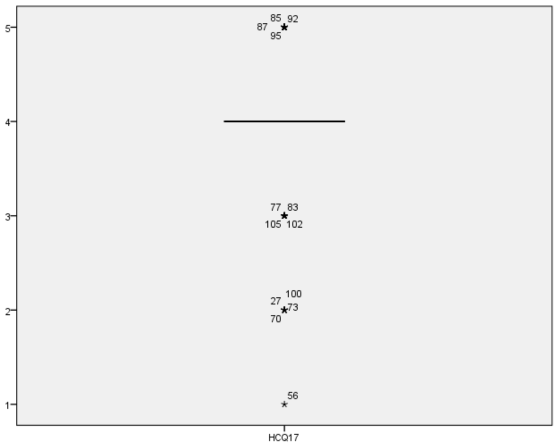
If you have any queries about this study please do not hesitate to contact Phil Birch (p.birch@chi.ac.uk) or Dr. Iain Greenlees (i.greenlees@chi.ac.uk).

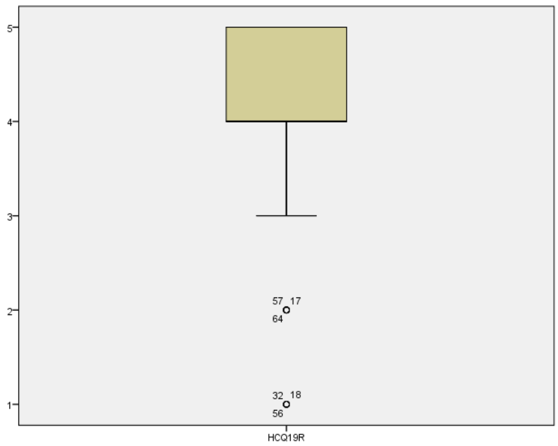
Thank you again for your participation

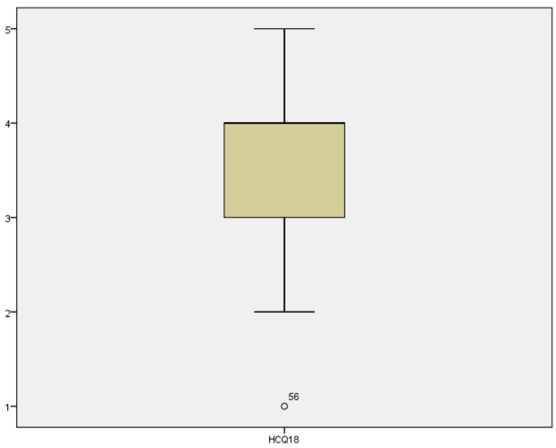
Appendix 3.4
Exploratory data: Individual boxplots for HCQ scores (data set A)

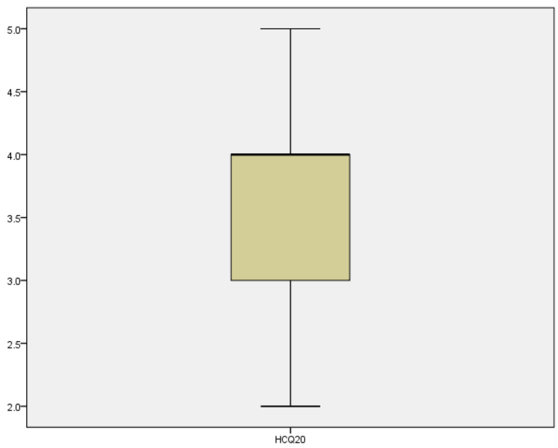


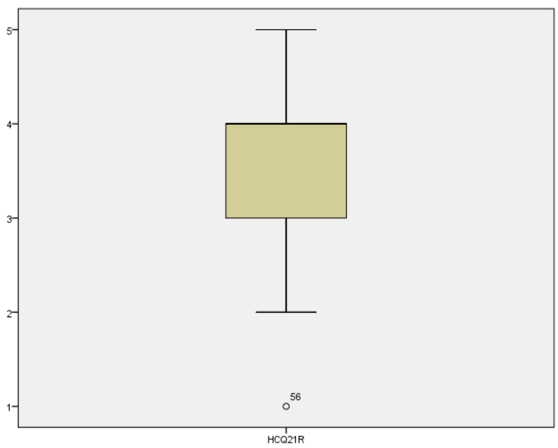

HCQ9R
HCQ11R
87 29
HCQ10
99 101
103 98
100 87
95
83 56
34
HCQ12
HCQ13
29 95 56
27
HCQ15R
100 56
57
HCQ14R
HCQ16
79 85
89 78
104 88
94
81 49
29

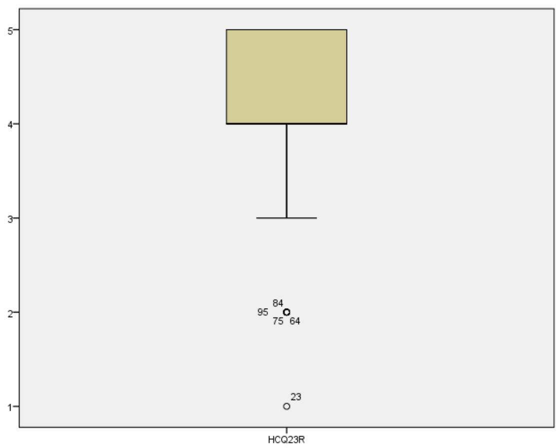

85 92
87
95
77 83
105 102
100
27
73
70
56
HCQ17

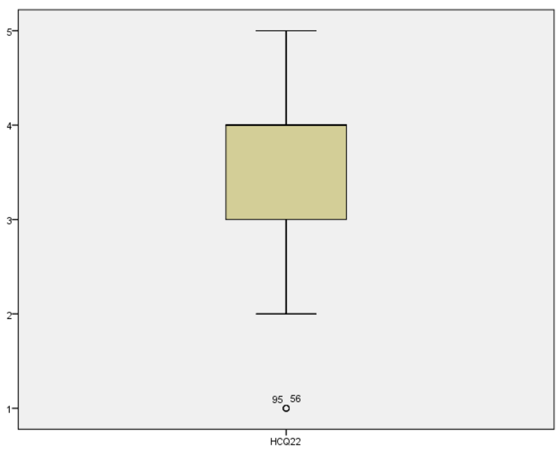

57 17
64
32 18
56
HCQ19R

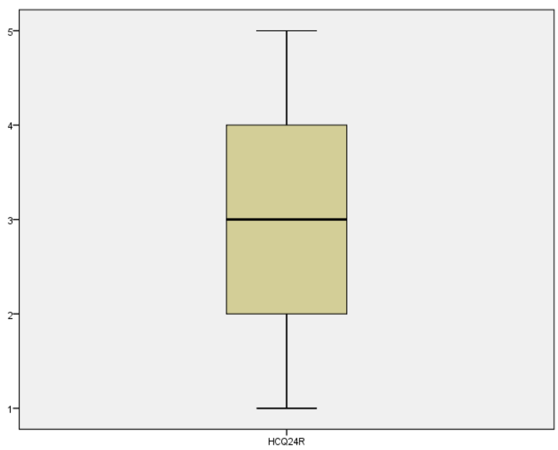

56
HCQ18

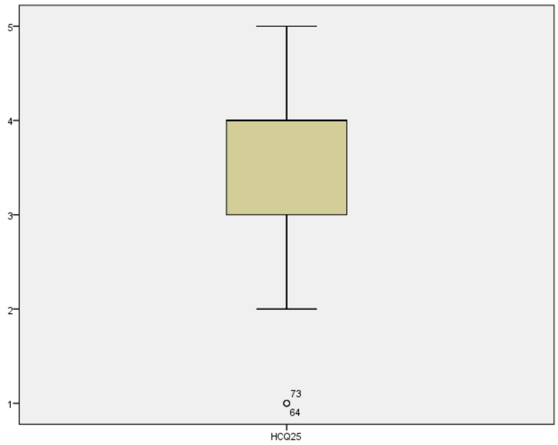

HCQ20

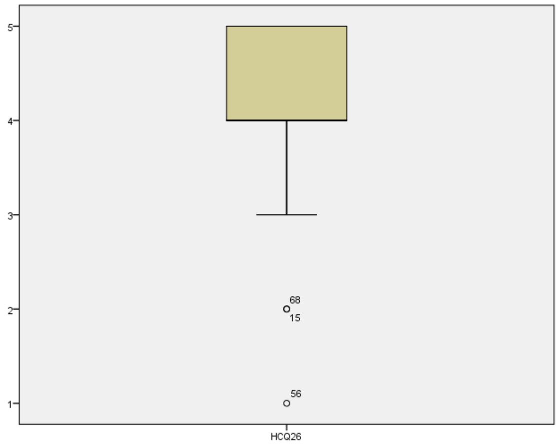

56
HCQ21R

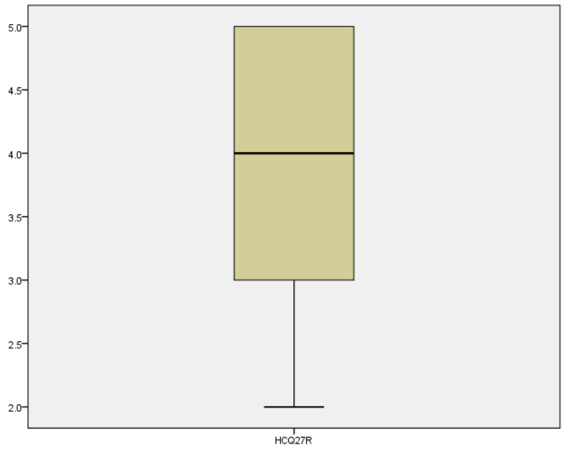

84
95
75 64
23
HCQ23R


95 56
HCQ22


HCQ24R


5
4
3
2
1
73
64
HCQ25


5
4
3
2
1
68
15
56
HCQ26

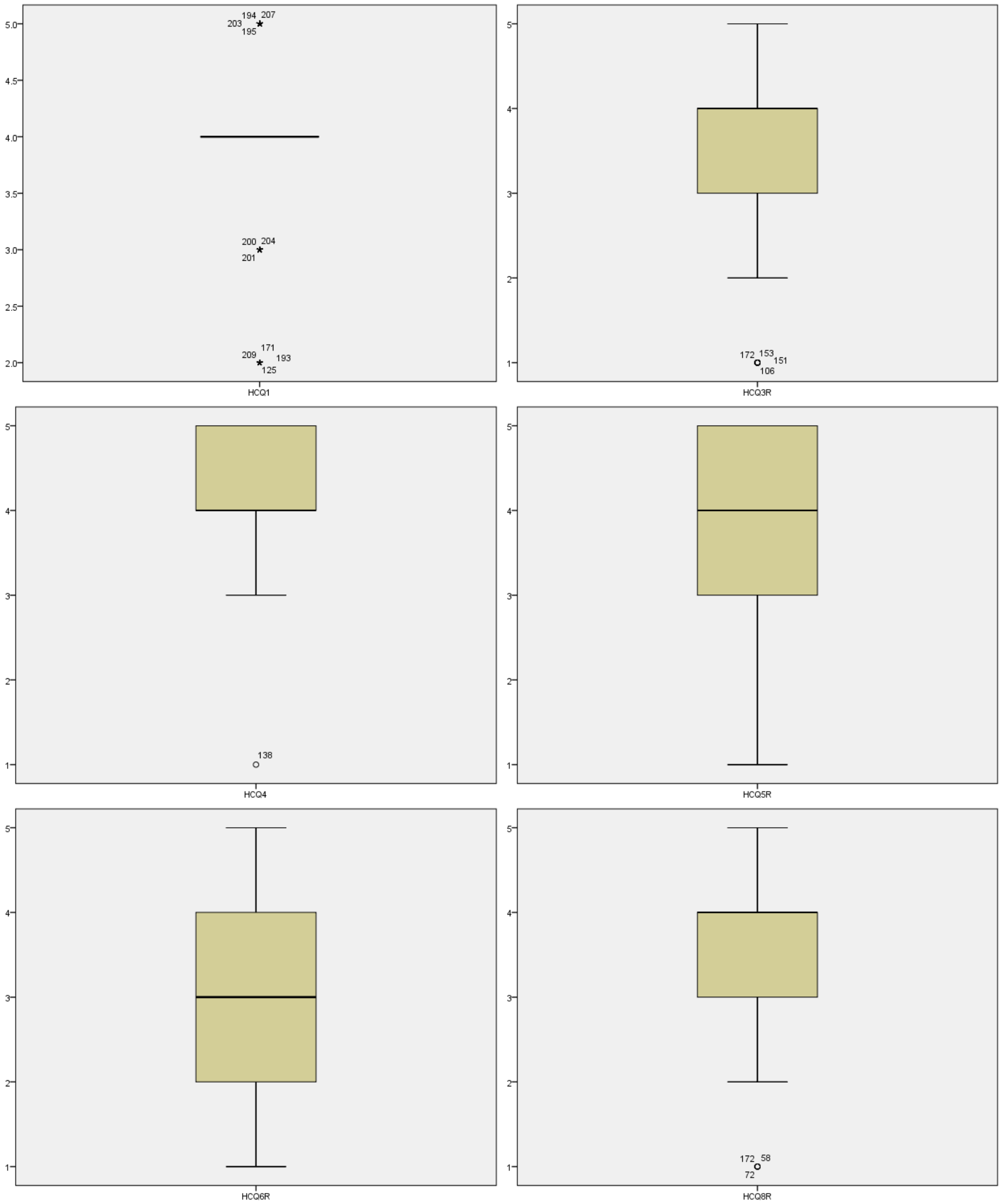

5.0
4.5
4.0
3.5
3.0
2.5
2.0
HCQ27R

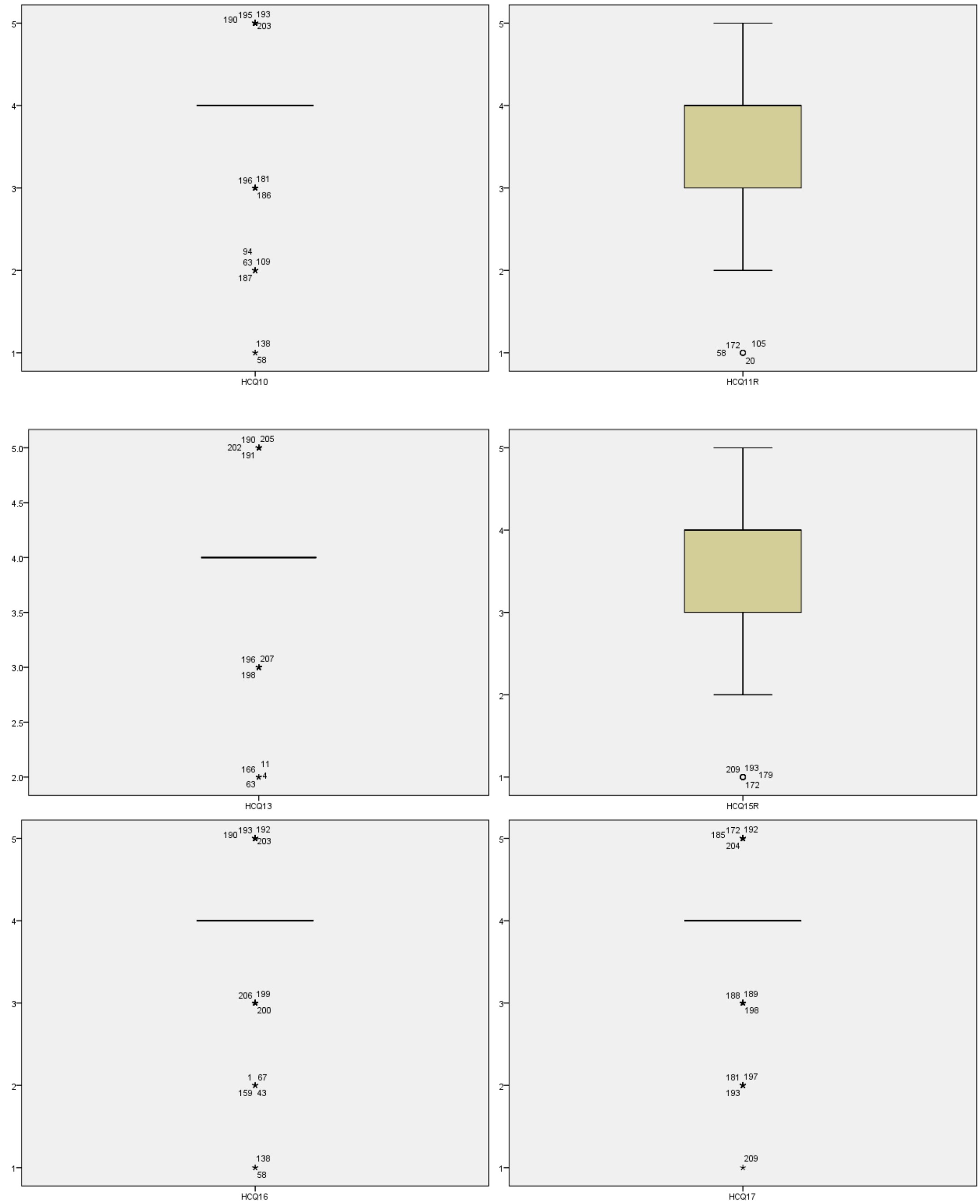
Appendix 3.5

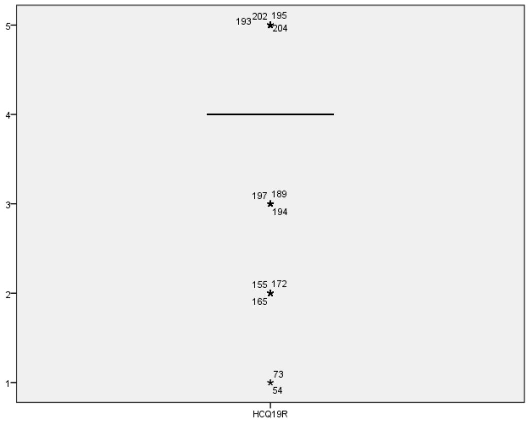
Exploratory data: Univariate normality statistics - extreme outliers removed (data set B)

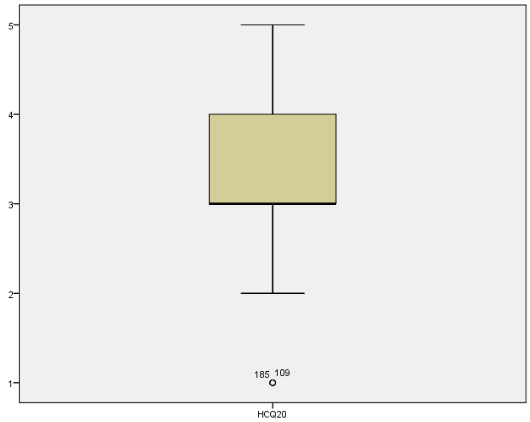
Item	Skewness ratio	Kurtosis ratio
HCQ1	-4.53	14.31
HCQ2	-1.81	5.10
HCQ3R	-2.16	-0.06
HCQ4	-0.91	-1.06
HCQ5R	-2.39	-0.19
HCQ6R	-0.39	-1.17
HCQ7	0.50	2.85
HCQ8R	-2.64	0.11
HCQ9R	-1.11	-1.20
HCQ10	0.04	-0.80
HCQ11R	-2.49	0.66
HCQ12	-1.69	0.55
HCQ13	-1.53	0.50
HCQ14R	-0.79	-1.13
HCQ15R	-1.40	-1.37
HCQ16	-0.56	-1.28
HCQ17	-1.61	0.86
HCQ18	-2.04	1.40
HCQ19R	-5.84	5.20
HCQ20	-1.22	0.07
HCQ21R	-2.10	-1.22
HCQ22	-1.37	-0.56
HCQ23R	-3.02	0.40
HCQ24R	-0.06	-0.40
HCQ25	-3.14	2.92
HCQ26	-2.78	2.50
HCQ27R	-2.29	-0.55

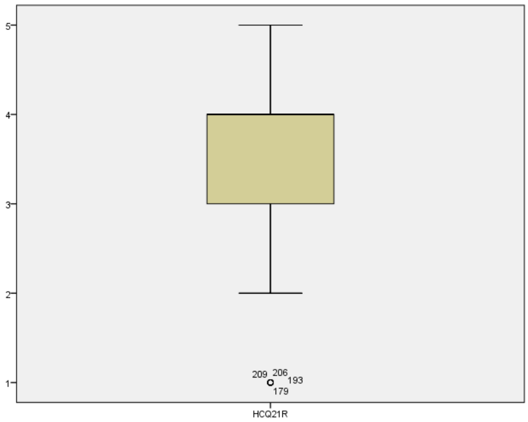
Appendix 3.6
Confirmatory data: Individual boxplots for HCQ scores (data set A)

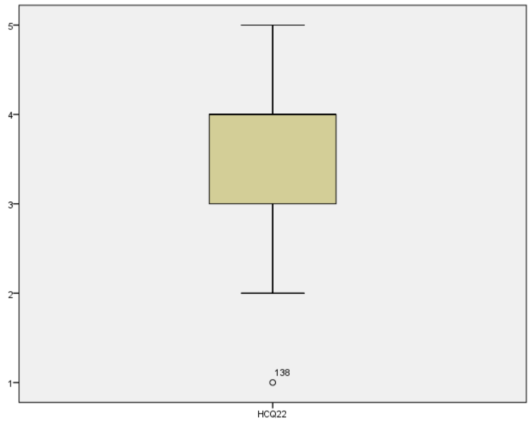


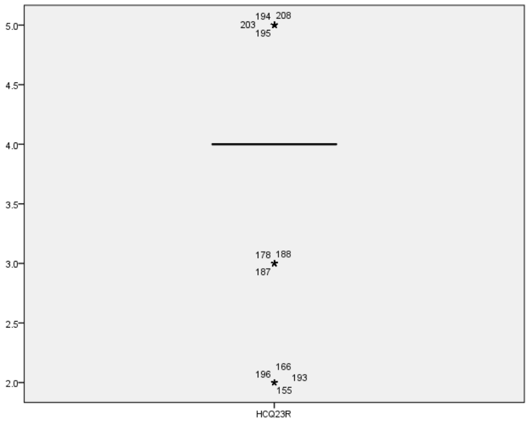

HCQ10
HCQ11R
HCQ13
HCQ15R
HCQ16
HCQ17

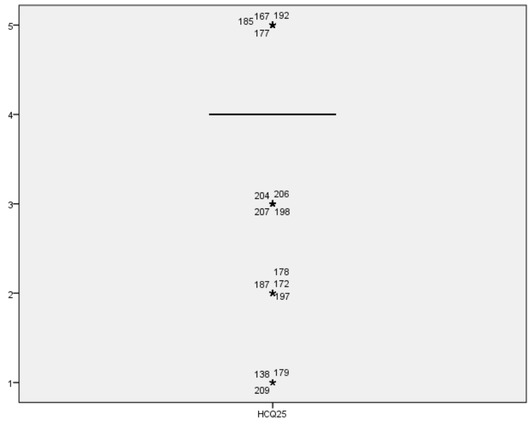

HCQ19R


HCQ20


HCQ21R


HCQ22


HCQ23R


HCQ25

Appendix 3.7
Confirmatory data: Multivariate outliers (Mahalanobis distance)

Observation number	Mahalanobis d-squared	p1	p2
58	75.045	.000	.000
172	74.135	.000	.000
138	58.287	.000	.000
73	52.414	.000	.000
20	52.059	.000	.000
50	50.599	.000	.000
22	47.211	.000	.000
72	45.228	.000	.000
11	43.333	.001	.000
187	40.839	.002	.000
36	39.285	.003	.000
193	39.121	.003	.000
105	38.136	.004	.000
12	37.908	.004	.000
166	36.500	.006	.000
63	34.866	.010	.000
95	33.037	.017	.000
54	32.798	.018	.000
67	32.546	.019	.000
133	32.209	.021	.000
179	32.031	.022	.000
14	30.925	.029	.000
196	30.296	.035	.000
80	30.238	.035	.000
27	30.210	.035	.000
6	29.224	.046	.000
207	28.837	.050	.000
171	28.828	.051	.000
94	28.720	.052	.000
41	28.360	.057	.000
165	27.739	.066	.000
90	27.627	.068	.000
126	27.308	.073	.000
81	27.247	.074	.000
8	26.467	.090	.000
151	26.430	.090	.000
42	26.107	.097	.000
181	26.100	.097	.000
107	25.955	.101	.000
35	25.854	.103	.000
47	25.476	.112	.000
120	25.149	.121	.001
109	24.886	.128	.001

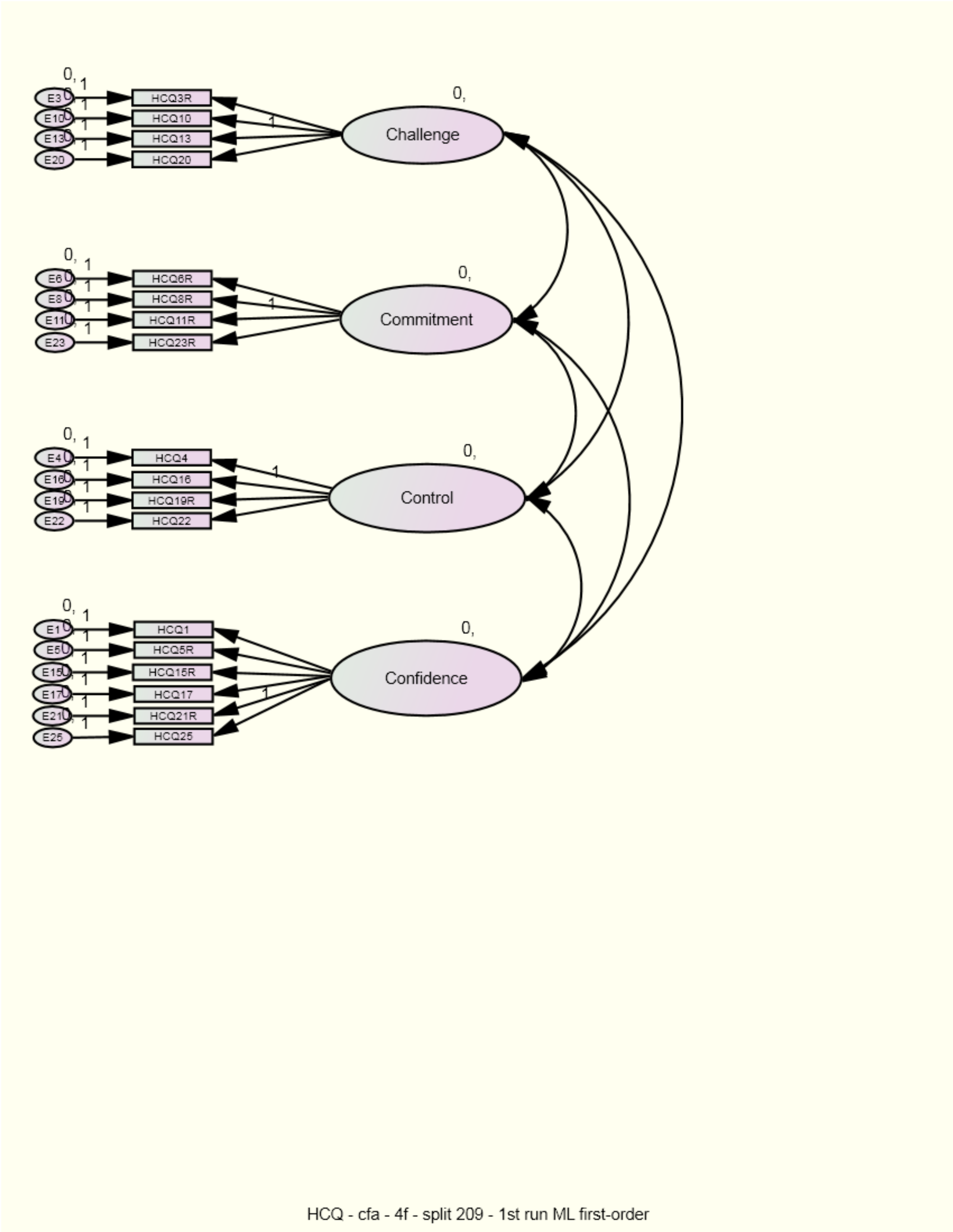
Observation number	Mahalanobis d-squared	p1	p2
188	24.745	.132	.001
4	24.374	.143	.003
104	24.088	.152	.006
155	24.045	.154	.004
96	23.943	.157	.004
209	23.823	.161	.004
124	23.691	.165	.004
43	23.513	.172	.005
116	23.461	.173	.004
136	23.019	.190	.014
123	22.877	.195	.016
21	22.704	.202	.020
137	22.696	.203	.014
185	22.085	.228	.076
24	21.761	.243	.138
204	21.704	.245	.124
162	21.607	.250	.124
144	21.268	.266	.221
202	21.212	.269	.203
106	21.135	.273	.195
1	21.042	.277	.195
127	20.999	.279	.173
150	20.778	.291	.234
97	20.610	.300	.276
206	20.463	.307	.310
78	20.214	.321	.412
178	20.179	.323	.379
64	20.038	.331	.416
92	19.790	.345	.529
46	19.660	.352	.562
2	19.288	.374	.749
28	19.113	.385	.800
122	18.588	.418	.952
82	18.521	.422	.950
129	18.464	.425	.946
70	18.282	.437	.964
186	18.227	.441	.961
84	18.095	.449	.970
134	17.827	.467	.988
153	17.647	.479	.993
149	17.594	.483	.992
39	17.557	.485	.991
88	17.447	.493	.992
195	17.443	.493	.989
183	17.401	.496	.987
31	17.389	.497	.983

Observation number	Mahalanobis d-squared	p1	p2
87	17.344	.500	.981
7	17.213	.509	.986
32	17.066	.519	.990
173	17.002	.523	.990
169	16.961	.526	.988
111	16.733	.542	.995
93	16.004	.592	1.000
121	15.888	.600	1.000
53	15.659	.616	1.000
23	15.619	.619	1.000
100	15.471	.629	1.000

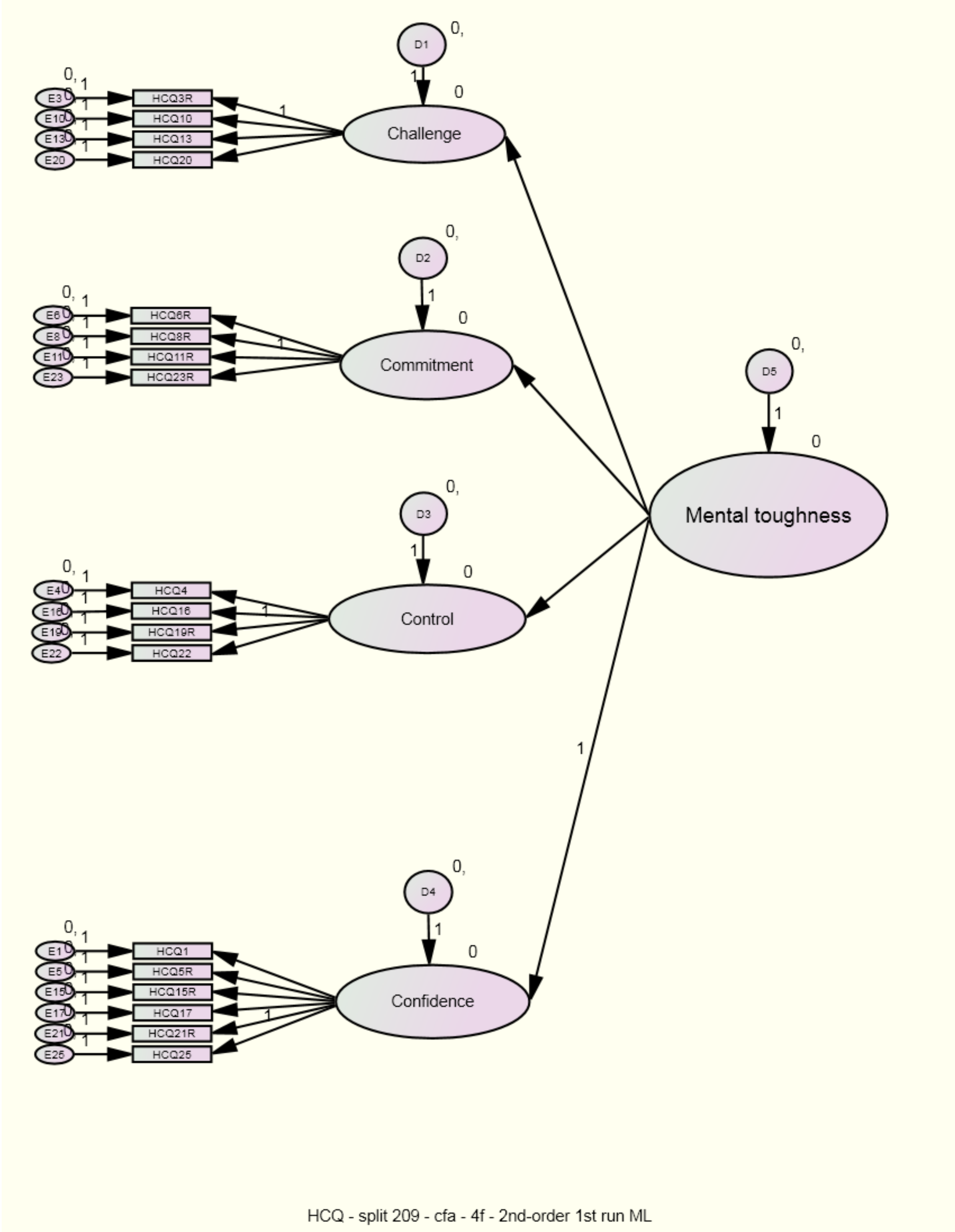
Appendix 3.8
Confirmatory data: Univariate normality statistics - extreme outliers removed (data set B)

Item	Skewness ratio	Kurtosis ratio
HCQ1	-4.74	6.53
HCQ3R	-4.79	1.94
HCQ4	0.26	-0.32
HCQ5R	-4.77	1.78
HCQ6R	-0.32	-2.06
HCQ8R	-4.29	1.97
HCQ10	-2.34	1.87
HCQ11R	-4.90	3.85
HCQ13	0.16	0.10
HCQ15R	-3.97	0.02
HCQ16	0.06	-0.92
HCQ17	-0.90	1.12
HCQ19R	-4.95	4.06
HCQ20	-2.41	0.16
HCQ21R	-4.20	1.28
HCQ22	-3.30	1.66
HCQ23R	-3.63	2.44
HCQ25	-2.16	2.13

Appendix 3.9
Path diagrams of the respective specified models of the HCQ



AMOS path diagram depicting the first-order four factor model of the HCQ.



AMOS path diagram depicting the second-order four factor model of the HCQ.

Appendix 3.10
Initial Exploratory Factor Analysis eigenvalues

Total Variance Explained

Initial Eigenvalues			Extraction Sums of Squared Loadings			Rotation Sums of Squared Loadings		
Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
6.963	25.787	25.787	6.963	25.787	25.787	3.923	14.528	14.528
2.638	9.772	35.559	2.638	9.772	35.559	2.956	10.949	25.478
2.187	8.101	43.660	2.187	8.101	43.660	2.207	8.175	33.653
1.641	6.079	49.739	1.641	6.079	49.739	2.060	7.631	41.284
1.307	4.843	54.582	1.307	4.843	54.582	2.017	7.470	48.753
1.179	4.368	58.949	1.179	4.368	58.949	1.745	6.463	55.217
1.053	3.902	62.851	1.053	3.902	62.851	1.700	6.297	61.514
1.003	3.714	66.566	1.003	3.714	66.566	1.364	5.052	66.566
.965	3.573	70.138						
.909	3.367	73.505						
.801	2.968	76.473						
.713	2.640	79.113						
.660	2.445	81.559						
.652	2.416	83.974						
.577	2.136	86.111						
.544	2.016	88.127						
.487	1.803	89.930						
.432	1.598	91.528						
.373	1.382	92.910						

20	.349	1.292	94.202						
21	.330	1.223	95.425						
22	.308	1.140	96.565						
23	.248	.920	97.485						
24	.226	.837	98.322						
25	.192	.710	99.032						
26	.152	.562	99.594						
27	.110	.406	100.000						

Extraction Method: Principal Component Analysis.

Appendix 3.11
Parallel Analysis output

PA ANALYSIS

(http://www.stattools.net/Parallel_Pgm.php)

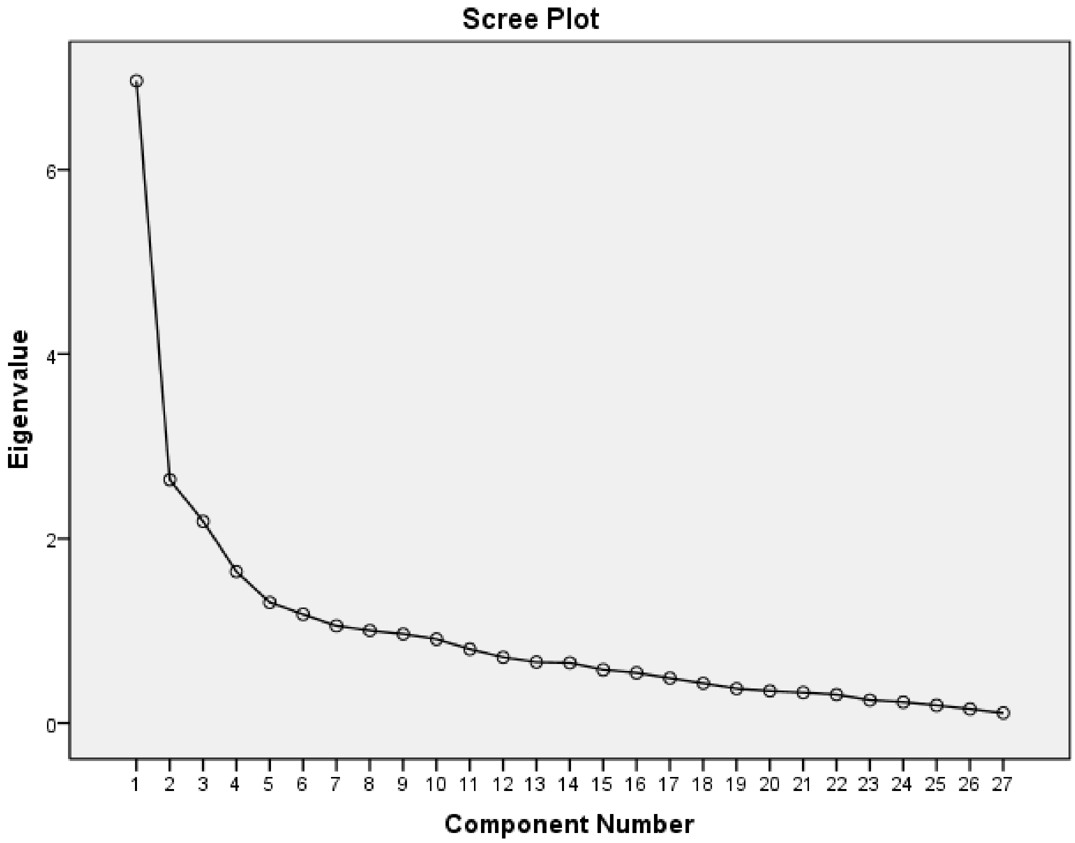
Monte Carlo Averaged Eigen Values

Averaged Variances (eigen values)
using normally distributed random numbers
Number of variables = 27
Sampe size of data = 105
Number of replications in simulation = 100

	Monte Carlo		
	Average	SD	95 percentile
Comp 1	2.0874	0.1025	2.2566
Comp 2	1.8961	0.0792	2.0268
Comp 3	1.7694	0.0562	1.8622
Comp 4	1.6666	0.0512	1.751
Comp 5	1.5681	0.0474	1.6463
Comp 6	1.4729	0.0445	1.5463
Comp 7	1.3826	0.0386	1.4463
Comp 8	1.3059	0.0394	1.371
Comp 9	1.2343	0.0388	1.2982
Comp 10	1.1592	0.0361	1.2188
Comp 11	1.0958	0.0354	1.1542
Comp 12	1.0349	0.0345	1.0918
Comp 13	0.9763	0.0319	1.029
Comp 14	0.9208	0.0309	0.9718
Comp 15	0.8679	0.0299	0.9174

Comp 16	0.8157	0.0317	0.868
Comp 17	0.7632	0.0311	0.8145
Comp 18	0.7123	0.0271	0.7569
Comp 19	0.6585	0.0286	0.7057
Comp 20	0.6133	0.0271	0.6581
Comp 21	0.5659	0.0288	0.6133
Comp 22	0.5225	0.026	0.5653
Comp 23	0.4758	0.0256	0.5181
Comp 24	0.4288	0.0224	0.4658
Comp 25	0.3844	0.0258	0.427
Comp 26	0.3376	0.0257	0.3799
Comp 27	0.2836	0.0305	0.3339

Appendix 3.12
Initial Exploratory Factor Analysis scree plot



Appendix 3.13
Initial Exploratory Factor Analysis rotated factor loadings

Rotated Component Matrix[a]

	Component							
	1	2	3	4	5	6	7	8
HCQ1	.684							
HCQ2			.621					
HCQ3R						.785		
HCQ4			.452		.428			
HCQ5R	.624							
HCQ6R		.777						
HCQ7				.743				
HCQ8R		.771						
HCQ9R								.648
HCQ10			.558			.464		
HCQ11R		.655						
HCQ12	.529							
HCQ13			.450	.645				
HCQ14R							.731	
HCQ15R	.595					.419		
HCQ16					.653			
HCQ17	.831							
HCQ18				.499				
HCQ19R					.555			
HCQ20			.742					
HCQ21R	.581					.458		
HCQ22					.723			
HCQ23R		.764						
HCQ24R							.436	
HCQ25	.884							
HCQ26								-.625
HCQ27R							.667	

Extraction Method: Principal Component Analysis.
Rotation Method: Varimax with Kaiser Normalization.
a. Rotation converged in 11 iterations.

Appendix 3.14
Secondary Exploratory Factor Analysis (fixed to four factors) eigenvalues

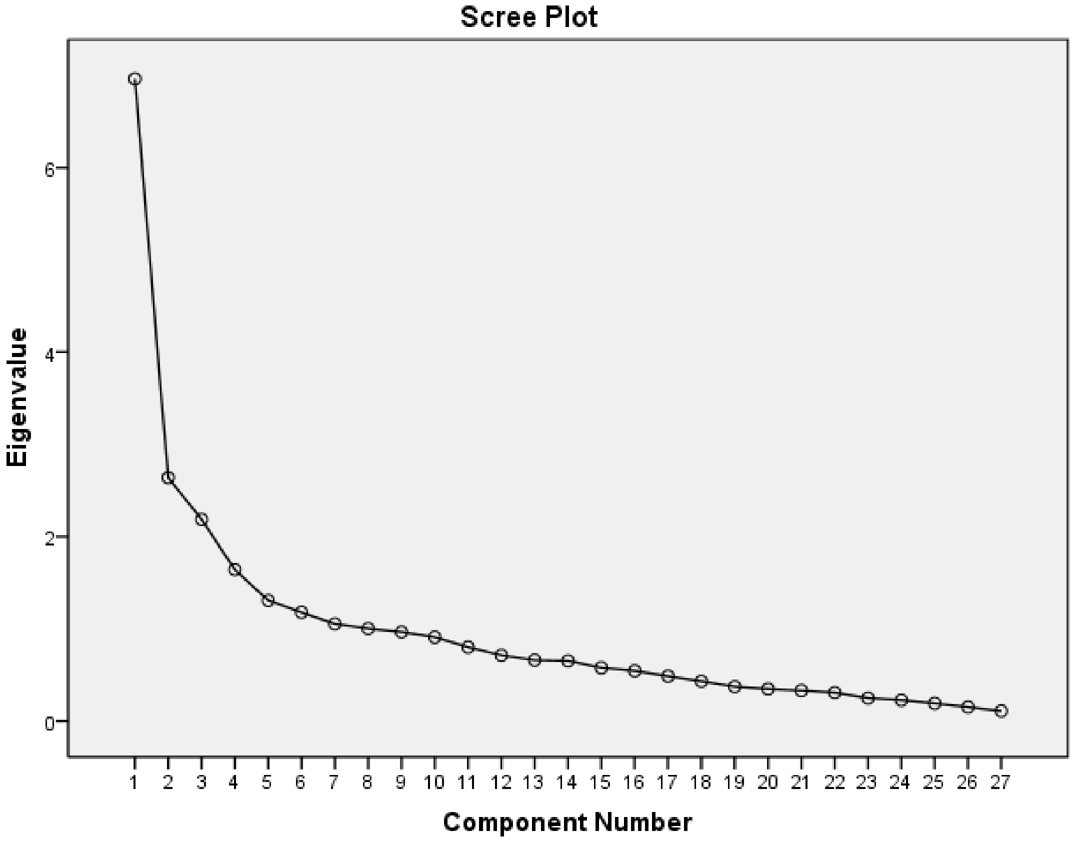
Total Variance Explained

Initial Eigenvalues			Extraction Sums of Squared Loadings			Rotation Sums of Squared Loadings		
Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
6.963	25.787	25.787	6.963	25.787	25.787	4.040	14.963	14.963
2.638	9.772	35.559	2.638	9.772	35.559	3.614	13.385	28.349
2.187	8.101	43.660	2.187	8.101	43.660	3.013	11.157	39.506
1.641	6.079	49.739	1.641	6.079	49.739	2.763	10.233	49.739
1.307	4.843	54.582						
1.179	4.368	58.949						
1.053	3.902	62.851						
1.003	3.714	66.566						
.965	3.573	70.138						
.909	3.367	73.505						
.801	2.968	76.473						
.713	2.640	79.113						
.660	2.445	81.559						
.652	2.416	83.974						
.577	2.136	86.111						
.544	2.016	88.127						
.487	1.803	89.930						
.432	1.598	91.528						
.373	1.382	92.910						

20	.349	1.292	94.202						
21	.330	1.223	95.425						
22	.308	1.140	96.565						
23	.248	.920	97.485						
24	.226	.837	98.322						
25	.192	.710	99.032						
26	.152	.562	99.594						
27	.110	.406	100.000						

Extraction Method: Principal Component Analysis.

Appendix 3.15
Secondary Exploratory Factor Analysis (fixed to four factors) scree plot



Appendix 3.16
Secondary EFA (fixed to four factors) rotated factor loadings

Rotated Component Matrix[a]

	Component			
	1	2	3	4
HCQ1	.740			
HCQ2				.549
HCQ3R				.433
HCQ4			.596	
HCQ5R	.597			
HCQ6R		.714		
HCQ7			.441	
HCQ8R		.725		
HCQ9R		.555		
HCQ10				.768
HCQ11R		.767		
HCQ12	.516			
HCQ13				.695
HCQ14R		.441		
HCQ15R	.651			
HCQ16			.546	
HCQ17	.783			
HCQ18			.455	
HCQ19R			.638	
HCQ20				.450
HCQ21R	.643			
HCQ22			.636	
HCQ23R		.723		
HCQ24R	.479			
HCQ25	.784			
HCQ26			.436	.467
HCQ27R		.496	.463	

Extraction Method: Principal Component Analysis.
Rotation Method: Varimax with Kaiser Normalization.
a. Rotation converged in 6 iterations.

Appendix 3.17
Confirmatory sample: Regression weights of the respective models of the HCQ

Standardised regression weights of the first-order four factor model of the HCQ

Item / factor			Estimate
HCQ10	<---	Challenge	0.711*
HCQ13	<---	Challenge	0.520*
HCQ20	<---	Challenge	0.367*
HCQ3R	<---	Challenge	0.336*
HCQ23R	<---	Commitment	0.751*
HCQ11R	<---	Commitment	0.734*
HCQ8R	<---	Commitment	0.730*
HCQ6R	<---	Commitment	0.658*
HCQ16	<---	Control	0.692*
HCQ4	<---	Control	0.678*
HCQ19R	<---	Control	0.467*
HCQ22	<---	Control	0.441*
HCQ25	<---	Confidence	0.777*
HCQ21R	<---	Confidence	0.751*
HCQ15R	<---	Confidence	0.736*
HCQ5R	<---	Confidence	0.698*
HCQ17	<---	Confidence	0.686*
HCQ1	<---	Confidence	0.542*

Note: * Indicates significant relationship ($p < 0.05$).

Standardised regression weights of the second-order four factor model of the HCQ

Item / factor			Estimate
Challenge	<---	Mental toughness	0.805*
Confidence	<---	Mental toughness	0.724*
Commitment	<---	Mental toughness	0.636*
Control	<---	Mental toughness	0.630*
HCQ10	<---	Challenge	0.711*
HCQ13	<---	Challenge	0.520*
HCQ20	<---	Challenge	0.367*
HCQ3R	<---	Challenge	0.336*
HCQ23R	<---	Commitment	0.751*
HCQ11R	<---	Commitment	0.734*
HCQ8R	<---	Commitment	0.730*
HCQ6R	<---	Commitment	0.658*
HCQ16	<---	Control	0.692*
HCQ4	<---	Control	0.678*
HCQ19R	<---	Control	0.467*
HCQ22	<---	Control	0.441*
HCQ25	<---	Confidence	0.777*
HCQ21R	<---	Confidence	0.751*
HCQ15R	<---	Confidence	0.736*
HCQ5R	<---	Confidence	0.698*
HCQ17	<---	Confidence	0.686*
HCQ1	<---	Confidence	0.542*

Note: * Indicates significant relationship ($p < 0.05$).

Appendix 3.18
Confirmatory sample: Inter-factor correlations of the revised first-order four factor model of the HCQ.

Correlations between factors of the first-order four factor revised model of the HCQ using the confirmatory sample.

	MT factors			
	Challenge	Commitment	Control	Confidence
Challenge	1			
Commitment	.31**	1		
Control	.29**	.24**	1	
Confidence	.26**	.37**	.41**	1

Note: **Significance is at $p < 0.01$.

Study 4 Assessment Instruments and SPSS Output

Appendix 4.1
The Revised Hardiness Confidence Questionnaire

Considering how you are **generally**, please answer the following statements. Start each statement with the phrase “In my general life.” For example, the first statement should be read “In my general life, I can influence the path that my life takes.” Please answer **honestly**, we are looking for your initial response to each statement so try not to spend too much time on any one statement.

Please indicate your response by **circling one** of the numbers, which have the following meaning;

1 = strongly disagree; **2** = disagree; **3** = neither agree nor disagree; **4** = agree; **5** = strongly agree.

In my general life, …	Strongly disagree	Disagree	Neither agree or disagree	Agree	Strongly agree
1. I can influence the path that my life takes	1	2	3	4	5
2. I find it difficult to stay committed to whatever I am doing	1	2	3	4	5
3. I doubt whether I am a worthwhile person	1	2	3	4	5
4. I believe that change enables me to grow as a person	1	2	3	4	5
5. I feel that I am a person of worth	1	2	3	4	5
6. Changes in my daily routine encourage me to learn	1	2	3	4	5
7. My actions have little influence on my life	1	2	3	4	5
8. I lack commitment	1	2	3	4	5
9. I feel that I am a truly worthwhile person	1	2	3	4	5
10. I find it difficult to stay committed to achieving my goals	1	2	3	4	5
11. I see challenges in my life as opportunities for me to develop as a person	1	2	3	4	5
12. I lack self-belief	1	2	3	4	5
13. Events in my life are shaped by my own actions	1	2	3	4	5
14. I take a negative attitude toward myself	1	2	3	4	5

Appendix 4.2
The Revised Hardiness Confidence Questionnaire scoring key

HCQ-R Scoring Key

Constructs	Items	Total number of items
Challenge	4 (10), 6 (20), 11 (13).	3 (0 X R)
Commitment	2R (8R), 8R (23R), 10R (11R)	3 (3 X R)
Control	1 (4), 7R (19R), 13 (16)	3 (1 X R)
Confidence	3R (5R), 5 (17), 9 (25), 12R (15R), 14R (21R)	5 (3 X R)

Note: Parentheses denotes HCQ item number

Calculating scores

Total mean MT = Add all 14 items scores (accounting for 7 reversed items) and divide by 14

Construct scores = Add the respective items scores (accounting for reversed items!) for each construct and divide by the number of items for that construct.

Appendix 4.3
The Revised Dispositional Resilience Scale-15

DRS-15

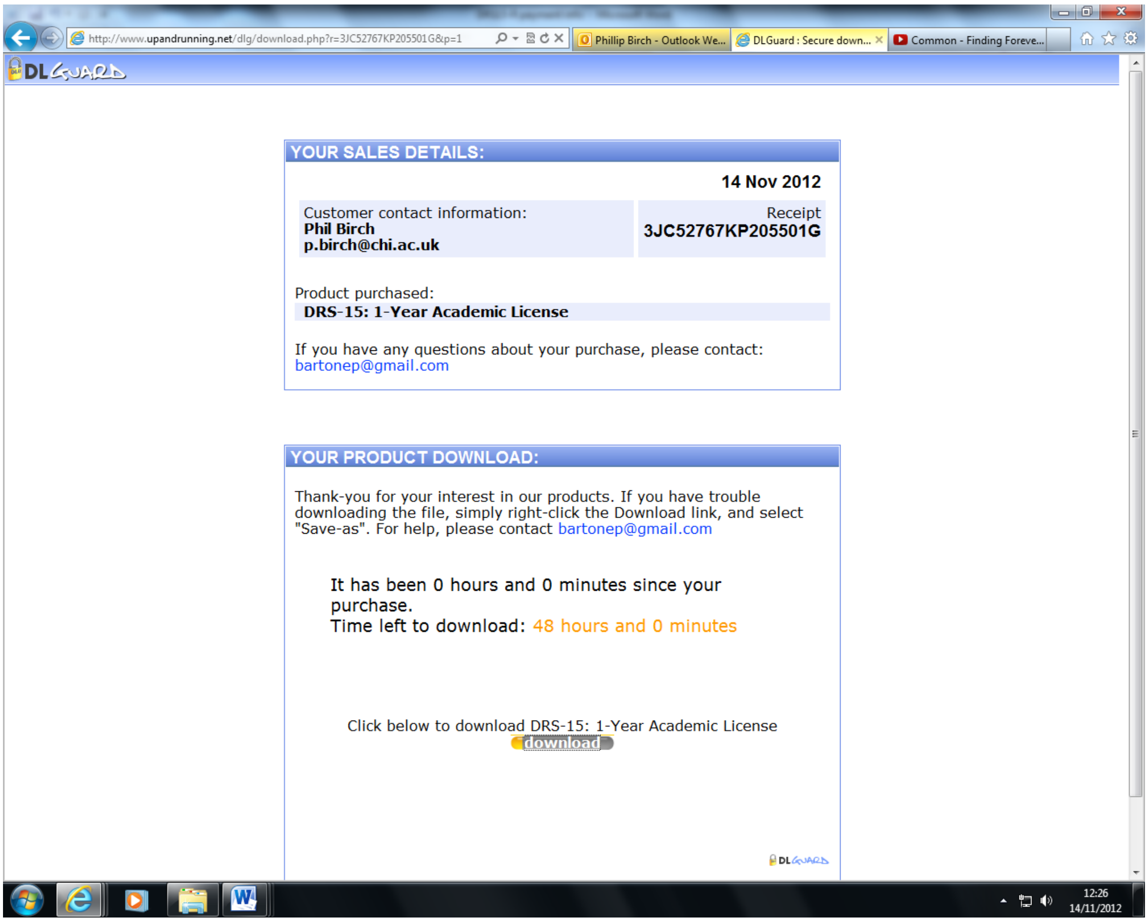
Below are statements about life that people often feel differently about.
Please show how much you think each one is true about you.
Give your own honest opinions . . . There are no right or wrong answers!

FILL IN THE BUBBLES TO SHOW YOUR ANSWERS

- Not at all true
- A little true
- Quite true
- Completely true

	Not at all true	A little true	Quite true	Completely true	Office use		
Most of my life gets spent doing things that are meaningful....................	◯	◯	◯	◯			
By working hard you can nearly always achieve your goals.........................	◯	◯	◯	◯			
I don't like to make changes in my regular activities..................................	◯	◯	◯	◯			
I feel that my life is somewhat empty of meaning..	◯	◯	◯	◯			
Changes in routine are interesting to me...	◯	◯	◯	◯			
How things go in my life depends on my own actions..................................	◯	◯	◯	◯			
I really look forward to my work activities...	◯	◯	◯	◯			
I don't think there is much I can do to influence my own future.............	◯	◯	◯	◯			
I enjoy the challenge when I have to do more than one thing at a time	◯	◯	◯	◯			
Most days, life is really interesting and exciting for me...........................	◯	◯	◯	◯			
It bothers me when my daily routine gets interrupted..............................	◯	◯	◯	◯			
It is up to me to decide how the rest of my life will be.............................	◯	◯	◯	◯			
Life in general is boring for me..	◯	◯	◯	◯			
I like having a daily schedule that doesn't change very much....................	◯	◯	◯	◯			
My choices make a real difference in how things turn out in the end.....	◯	◯	◯	◯			

Permission to use the Revised Dispositional Resilience Scale-15



Appendix 4.4
The Revised Dispositional Resilience Scale-15 scoring key

SCORING KEY FOR DRS-15 DISPOSITIONAL RESILIENCE SCALE (v.3)

	Below are statements about life that people often feel differently about. Please check a box to show how much you think each one is true for you. Give your own honest opinions… There are no right or wrong answers!		Not at all true	A little true	Quite true	Completely true
1	Most of my life gets spent doing things that are meaningful	**CM**	0	1	2	3
2	By working hard you can nearly always achieve your goals	**CO**	0	1	2	3
3	I don't like to make changes in my regular activities	**CH(-)**	3	2	1	0
4	I feel that my life is somewhat empty of meaning	**CM(-)**	3	2	1	0
5	Changes in routine are interesting to me	**CH**	0	1	2	3
6	How things go in my life depends on my own actions	**CO**	0	1	2	3
7	I really look forward to my work activities	**CM**	0	1	2	3
8	I don’t think there’s much I can do to influence my own future	**CO(-)**	3	2	1	0
9	I enjoy the challenge when I have to do more than one thing at a time	**CH**	0	1	2	3
10	Most days, life is really interesting and exciting for me	**CM**	0	1	2	3
11	It bothers me when my daily routine gets interrupted	**CH(-)**	3	2	1	0
12	It is up to me to decide how the rest of my life will be	**CO**	0	1	2	3
13	Life in general is boring for me	**CM(-)**	3	2	1	0
14	I like having a daily schedule that doesn't change very much	**CH(-)**	3	2	1	0
15	My choices make a real difference in how things turn out in the end	**CO**	0	1	2	3

SCORES ARE REVERSED ON SIX NEGATIVELY KEYED ITEMS: 3, 4, 8, 11, 13, 14

CM = COMMITMENT = SUM (1+4+7+10+13)

CO=CONTROL = SUM (2+6+8+12+15)

CH=CHALLENGE = SUM (3+5+9+11+14)

TOTAL HARDINESS-RESILIENCE SCORE = SUM (CM+CO+CH)

Appendix 4.5
The Revised Self-liking/self-competence Scale

Please indicate how much you agree with each of the 16 statements below.

Be as honest and as accurate as possible. Do not skip any statements.

Respond to the statements in the order they appear. Use the following scale:

1	2	3	4	5
strongly disagree				strongly agree

	Strongly disagree				Strongly agree
1. I tend to devalue myself	1	2	3	4	5
2. I am highly effective at the things I do	1	2	3	4	5
3. I am very comfortable with myself	1	2	3	4	5
4. I am almost always able to accomplish what I try for	1	2	3	4	5
5. I am secure in my sense of self-worth	1	2	3	4	5
6. It is sometimes unpleasant for me to think about myself	1	2	3	4	5
7. I have a negative attitude toward myself	1	2	3	4	5
8. At times, I find it difficult to achieve the things that are important to me	1	2	3	4	5
9. I feel great about who I am	1	2	3	4	5
10. I sometimes deal poorly with challenges	1	2	3	4	5
11. I never doubt my personal worth	1	2	3	4	5
12. I perform very well at many things	1	2	3	4	5
13. I sometimes fail to fulfil my goals	1	2	3	4	5
14. I am very talented	1	2	3	4	5
15. I do not have enough respect for myself	1	2	3	4	5
16. I wish I were more skilful in my activities	1	2	3	4	5

Appendix 4.6
The Revised Self-liking/Self-competence Scale scoring key

SLSC-R Scoring Key:

SL = sum of 3, 5, 9, 11, 1*, 6*, 7*, 15*

SC = sum of 2, 4, 12, 14, 8*, 10*, 13*, 16*
*reverse-scored (6 minus score)

Add the respective items scores (accounting for reversed items) for each construct and divide by 8.

Appendix 4.7
The Coping Function Questionnaire

Describe the most stressful situation in sport that YOU have faced in the last 12 months?

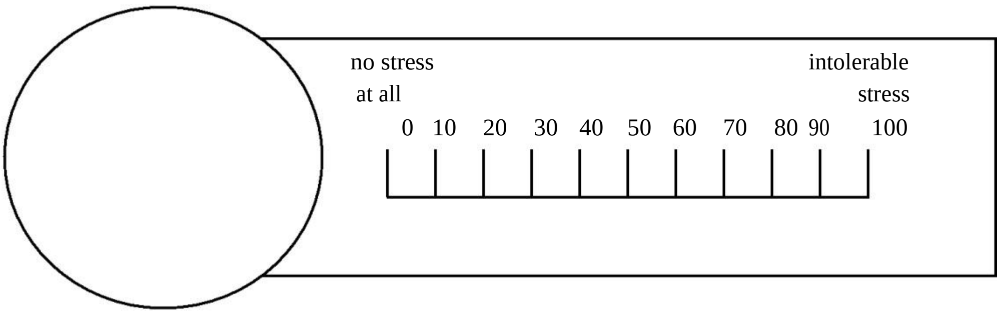
How long did the stressful situation last? (Circle one)

less than 1 week   1 week to 1 month   1 to 3 months   more than 3 months

When did the stressful situation occur? (Circle one)

in the past week	more than 1 week ago, but less than 1 month ago	more than 1 month ago, but less than 3 months ago	more than 3 months ago, but less than 12 months ago

Please indicate the amount of stress that you experienced in the situation by marking an 'X' on the scale within the thermometer:



Why was this situation stressful to you?

We are interested in how you tried to handle the stressful situation you described above. We are NOT concerned with what
you did before the situation happened or what you did after it was over.
We want to know what you did to try to handle the stressful situation when you were actually faced with it.

****Take a minute to think about the types of things you did to try to handle the situation and then continue.****

We would like you to indicate, by filling in the appropriate circle on the next page, to what extent you used each of the following to deal with the stress.

Remember:
A. There are no right and wrong answers - this is not a test.

B. Please answer all the questions as honestly and accurately as you can - this is very important.

How much did you use the following to deal with the stressful situation you described on the previous page?

1. **I tried to find a way to change the situation**.

Not at all | A little | Somewhat | Quite a bit | Very much

2. **I stayed in the situation and tried to control my emotions to better deal with the situation**.

Not at all | A little | Somewhat | Quite a bit | Very much

3. **I worked harder to try to change the situation**.

Not at all | A little | Somewhat | Quite a bit | Very much

4. **I tried to change how I thought about the situation so it didn't seem so stressful**.

Not at all | A little | Somewhat | Quite a bit | Very much

5. **I tried to get out of the situation as soon as I could to reduce the stress**.

Not at all | A little | Somewhat | Quite a bit | Very much

6. **I used strategies to change the situation in order to deal with the stress**.

Not at all | A little | Somewhat | Quite a bit | Very much

7.

I tried to view the situation in a way that made it seem less stressful.

Not at all A little Somewhat Quite a bit Very much

8.

I tried to leave or avoid the situation
to get away from the problem or reduce the stress.

Not at all A little Somewhat Quite a bit Very much

9.

I did my best to change the situation.

Not at all A little Somewhat Quite a bit Very much

10.

I tried to use
different strategies that
would help me control my
emotions.

Not at all A little Somewhat Quite a bit Very much

11.

I looked for ways to solve the problem or change the situation.

Not at all A little Somewhat Quite a bit Very much

12.

I tried to get out of the situation to get away from the stress.

Not at all A little Somewhat Quite a bit Very much

13.

I stayed in the situation and tried to change it.

Not at all A little Somewhat Quite a bit Very much

14.

I worked through my emotions in order to feel better.

Not at all A little Somewhat Quite a bit Very much

15.

I tried to get away from the situation to reduce the stress.

Not at all A little Somewhat Quite a bit Very much

16.

I tried to find ways to control my emotions.

Not at all A little Somewhat Quite a bit Very much

17.

I tried to relax so that I could keep my emotions under control.

Not at all A little Somewhat Quite a bit Very much

18.

In order to reduce the stress
I tried to get myself out of the situation.

Not at all A little Somewhat Quite a bit Very much

Appendix 4.8
The Coping Function Questionnaire scoring key

Introduction.

The Coping Function Questionnaire (CFQ) was developed to assess problem-focused, emotion-focused, and avoidance coping for adolescent sport participants. The CFQ assesses coping *function* as opposed to specific coping strategies. It has 18 items that assess problem-focused (six items), emotion- focused (seven items), and avoidance (five items) coping function. On the CFQ, participants are asked how much they used each coping function to handle a self-indicated stressful situation.

Problem-focused coping items are: 1,3,6,9,11,13
Emotion-focused coping items are: 2,4,7,10,14,16,17
Avoidance coping items are: 5,8,12,15,18

Responses for each coping function item are scored on a 5-point scale (not at all, a little, somewhat, quite a bit, very much). Coping function scale scores are determined by taking the mean of all items comprising each scale, with higher scores reflecting greater coping. Scale scores ranged from 1 (not used) to 5 (used very much).

CFQ development and validation information can be found in the following journal article:

Kowalski, K. C., & Crocker, P. R. E. (2001). Development and validation of the Coping Function Questionnaire for adolescents in sport. *Journal of Sport and Exercise Psychology*, *23*, 136-155.

****
***

Keys to Successful Administration of the CFQ

Read instructions on the first page emphasizing that they **"describe the most stressful situation in sport that YOU have faced in the last 12 months".**

Verbal instructions we have used also include:

> **Describe the most stressful situation you have experienced in the last 12 months. Just describe one (the most stressful) situation. This situation can be anything related to sport. It could be during competition, or it can be any issue surrounding your sport. If you have not participated in sport in the past 12 months, then you should describe a stressful situation you experienced in a physical activity setting, fitness/dance setting, physical education, or other type of game.**

Get them to fill out the questionnaire package, with a research assistant available to answer questions.

Once they are done, ask the participants to check over the questionnaire package just to make sure they haven't missed any of the items Explain to the participants that the research assistants are not looking at their coping behaviour, but rather just making sure they haven't missed any of the questions.

Limitations of the CFQ.

C. The CFQ is an appropriate option when researchers are interested in questions related to the *function* of coping strategies. A limitation of the CFQ is that does not provide information regarding what specific coping strategies are being used to achieve a function of coping. The CFQ provides measurement of coping at a global, functional level.

D. Although the three coping function structure on the CFQ (problem-focused, emotion-focused, avoidance) was developed from a theoretical framework, and the factor structure of the instrument supported this distinction (see Kowalski & Crocker, 2001), it is possible that there may be other functions coping serve that are not captured on the CFQ.

Appendix 4.9
Participant information sheet, consent form, demographic questionnaire, and written debrief

<u>Information Sheet for Participants</u>

<u>PLEASE READ THE FOLLOWING CAREFULLY</u>

Study title: **An examination of the psychological characteristics and thought processes of competitive athletes**

What is the study about?
The purpose of this study is to examine the psychological characteristics and thought processes of competitive athletes.

At the University of Chichester before any project starts it has to be checked by university staff. They make sure that the research is good enough to carry out. This project has been checked by the staff of the Department of Sport and Exercise Science and has been passed by the University Ethics Committee.

Do you have to take part?
No! It is your choice whether you participate in the project or not. If at any time during the project you feel that you don't want to continue then you can tell the researcher that you want to stop – you do not have to give a reason.

What will you do in the project?
The study involves completing a series of psychological questionnaires. You will be asked to complete two questionnaires relating to your general personality in the first sitting and four questionnaires relating to your general personality in this second sitting. Please take your time to complete the questionnaires and answer as truthfully as possible.

In total, the first batch of questionnaires should take no longer than 8 minutes and the second batch of questionnaires should take no longer than 12 minutes so the total amount of time you will need to commit to this project is 20 minutes.

Why have you been asked to take part?

You have been asked to take part in this project since we need athletes who are competing in sport to gain an insight into the psychology of sport.

What happens to the information in the project?
All the information which is collected about you during the project will be kept private – the data will be identifiable for data analysis purposes but will only be observed by the primary researcher. Only group data will be used for publication purposes which will make it impossible to trace this information back to you individually.

Thank you for your time - please ask any questions if you are unsure or confused about what we have said.

What happens if you agree to take part?
- If you are happy to be involved in the project then please sign the informed consent form and the accompanied questionnaires.
- If you do not want to be involved in the project then thank you.

Who can you contact if you have any questions about the project?

Phil Birch – Email p.birch@chi.ac.uk: Phone 01243 816343 or
Dr Iain Greenlees - Email i.greenlees@chi.ac.uk: Phone 01243 816437

Thank you for your time

Consent Form

I understand that my participation in this project will involve completing 6 questionnaires which will assess my general personality. I understand that I will be asked to complete two questionnaires today and the remaining 4 questionnaires in one week's time. I understand that participation in this study is entirely voluntary and that I can withdraw from the study at any time without giving a reason.

I understand that I am free to ask any questions at any time. I am free to withdraw or discuss my concerns with Phil Birch (p.birch@chi.ac.uk) or Dr. Iain Greenlees (i.greenlees@chi.ac.uk).

I understand that the information I provide over the two collection periods will be identifiable for data analysis purposes but will only be observed by the primary researcher. I understand that only group data will be used for publication purposes which will make it impossible to trace this information back to me individually. I understand that this information will be stored confidentially and may be retained indefinitely.

I also understand that at the end of the study I will be provided with additional information and feedback about the purpose of the study.

I, ____(NAME) consent to participate in the study conducted by Phil Birch

Signed:

Date:

Demographic Questionnaire

Please fill in your details in the spaces provided

Age ______________

(years)

Gender Male ☐ Female ☐

Ethnic origin – Please indicate the most appropriate description of your ethnic origin.

(Please tick)

White - British	☐	Asian - Indian	☐
White - Irish	☐	Asian - Pakistani	☐
White – Any other White background	☐	Asian – Bangladeshi	☐
Mixed – White & Black Caribbean	☐	Black or Black British - Caribbean	☐
Mixed – White & Black African	☐	Black or Black British - African	☐
Mixed – White & Black Asian	☐	Black or Black British - Any other Black background	☐
Mixed – Any other Mixed background	☐	Chinese – Any Chinese background	☐
Any other ethnic background	☐		

University of study ____

Year of study ____

(e.g., 1, 2, 3)

What is your primary sport? ____

How many years competitive playing experience do you have in your primary sport? ___________________________

What is the highest level that you have played in your primary sport? ____

Participant Debrief

The aim of this study was to examine the suitability of the Hardiness Confidence Questionnaire revised version (HCQ-R) in measuring mental toughness in sport and to devise programmes to enhance the suitability of the HCQ-R.

This required us to compare the HCQ-R with a number of other psychological questionnaires (i.e., the Coping Function Questionnaire, the Dispositional Resilience Scale, the Self-Liking/Self-Competence Scale, and the Social Desirability Scale).

If you have any queries about this study please do not hesitate to contact Phil Birch (p.birch@chi.ac.uk) or Dr. Iain Greenlees (i.greenlees@chi.ac.uk).

Thank you again for your participation

Appendix 4.10
Univariate normality statistics of the Revised Dispositional Resilience Scale-15, the Revised Self-liking/self-competence Scale, and the Coping Function Questionnaire

Univariate normality statistics for the Revised Dispositional Resilience Scale-15

Item	Skewness ratio	Kurtosis ratio
Challenge		
DRS15R1	-0.96	0.19
RDRS15R4	-6.94	4.78
DRS15R7	-0.17	-0.47
DRS15R10	-1.33	-0.31
RDRS15R13	-10.28	17.97
Commitment		
DRS15R2	-5.24	5.14
DRS15R6	-3.04	1.33
RDRS15R8	-14.29	28.11
DRS15R12	-2.67	-2.22
DRS15R15	-4.96	4.47
Control		
RDRS15R3	-1.60	0.00
DRS15R5	0.28	-0.70
DRS15R9	-0.10	-2.18
RDRS15R11	-2.78	-0.03
RDRS15R14	-3.63	1.53

Univariate normality statistics for the Revised Self-liking/self-competence Scale

Item	Skewness ratio	Kurtosis ratio
Self-liking		
RSCLSR1	-1.91	-1.64
SCLSR3	-2.31	-0.39
SCLSR5	-2.87	1.32
RSCLSR6	-1.98	-0.64
RSCLSR7	-2.00	-1.82
SCLSR9	-0.92	-0.75
SCLSR11	-0.78	-1.66
RSCLSR15	-0.56	-1.80
Self-competence		
SCLSR2	-0.25	-0.36
SCLSR4	-2.28	1.92
RSCLSR8	0.03	-1.80
RSCLSR10	-2.44	-0.08
SCLSR12	-1.76	1.38
RSCLSR13	-1.19	-1.59
SCLSR14	0.13	-0.06
RSCLSR16	2.38	-0.60

Univariate normality statistics for the Coping Function Questionnaire

Item	Skewness ratio	Kurtosis ratio
Problem-focused coping		
CFQ1	0.01	-2.80
CFQ3	-2.44	-1.72
CFQ6	0.68	-2.61
CFQ9	-1.28	-2.41
CFQ11	-1.06	-2.09
CFQ13	-1.65	-2.00
Emotional-focused coping		
CFQ2	-1.84	-1.74
CFQ4	-1.13	-1.95
CFQ7	-1.46	-2.11
CFQ10	-0.29	-2.53
CFQ14	-2.01	-0.55
CFQ16	-0.94	-1.64
CFQ17	-2.07	-1.22
Avoidance-focused coping		
CFQ5	3.56	-1.60
CFQ8	5.14	0.51
CFQ12	5.74	1.28
CFQ15	4.55	-0.37
CFQ18	4.57	-0.47

Appendix 4.11
The challenge and threat construal measure

Instructions: Please respond to the following statements in relation to how you feel about the upcoming competition. ***Some players would find such a game to be threatening to them*** *(i.e., they would see the game as being possibly harmful to their confidence, self-esteem).* ***Others would view this important match as a positive challenge*** *(i.e., they would see the game as a chance to* develop their skills, and play to the best of their abilities). *Please circle the answer that you feel is most true of you.*

Q	*How are you feeling in relation to today's match?*	Not at all true of me	Not true of me	Somewhat true of me			True of me	Very true of me
1	I view today's match as a positive challenge.	1	2	3	4	5	6	7
2	I view today's match as a threat	1	2	3	4	5	6	7
3	I am looking forward to being positively challenged in today's match	1	2	3	4	5	6	7
4	I think today's match could be threatening to me	1	2	3	4	5	6	7
5	I think today's match represents a positive challenge to me	1	2	3	4	5	6	7
6	I think today's match represents a threat to me	1	2	3	4	5	6	7

Appendix 4.12
The challenge and threat construal measure scoring key

<u>The challenge and threat construal measure scoring key</u>

Challenge: 1, 3, 5.

Threat: 2, 4, 6.

Add the respective items scores for each construct and divide by 3.

Appendix 4.13
The Immediate Anxiety Measurement Scale

Modern day sport, by nature is highly competitive, which is likely to cause stress and anxiety to those who participate. There are two main types of anxiety which may be experienced in a sports player. These are cognitive and somatic anxiety. The following questionnaire asks you to respond to how cognitively anxious you are (the mental component), how somatically anxious you are, (the physical component) and also how self-confident you are at this moment in time. In order to answer as accurately as possible please bear the following definitions in mind:

Cognitive Anxiety: Is the **mental** component of anxiety and maybe characterised by thoughts such as concerns or worries about your upcoming competition/match, for example about the way you will perform or the importance of the event.

Somatic Anxiety: Is your **perception** of your **physical** state and maybe characterised by symptoms such as physical nervousness, butterflies in the stomach, tense muscles, and increases in heart rate.

Self Confidence: Is how **confident** you are of performing well in your upcoming competition/match and maybe characterised by factors such as achieving your competition/match goals and performing well under pressure.

Section 1: 'Intensity': Refers to the **amount** or **level** of cognitive anxiety, somatic anxiety and self-confidence that you are experiencing right now.

Section 2: 'Directional perceptions': Refers to the extent to which you regard the intensity of these anxiety and confidence symptoms as positive or negative towards your upcoming performance.

Section 3: 'Frequency of intrusions': Refers to how **frequently** or **how often** you are experiencing these anxiety and confidence symptoms right now.

Instructions: Below are **3** statements reflecting the thoughts and feelings players may experience before a game. Each statement requires a response from each of the **3** sections. Section **1** asks you to respond to the **level** of cognitive anxiety, somatic anxiety, and self-confidence, (see definitions); Section **2** then asks whether you regard these anxiety and confidence level as **positive** or **negative** to your upcoming performance; Finally, Section **3** asks how **frequently** these anxiety and confidence symptoms are occurring at this time. Read each statement carefully and then circle the appropriate number to show how you feel **RIGHT NOW** in each of the 3 sections.

	SECTION 1	SECTION 2	SECTION 3
	To what extent are you experiencing the **anxiety** and **confidence**, (i.e. what **level**).	When you experience this **anxiety** and **confidence** do you regard it as **positive** or **negative** in relation to the upcoming competition/match.	How **frequently** are you experiencing this **anxiety** and **confidence.**
	Not at all … Extremely	Very Debilitative (Negative) … Unimportant … Very Facilitative (Positive)	Not at all … All of the time
Statement 1. **I am cognitively anxious**	1 2 3 4 5 6 7	-3 -2 -1 0 +1 +2 +3	1 2 3 4 5 6 7
Statement 2. **I am somatically anxious**	1 2 3 4 5 6 7	-3 -2 -1 0 +1 +2 +3	1 2 3 4 5 6 7
Statement 3. **I am self-confident**	1 2 3 4 5 6 7	-3 -2 -1 0 +1 +2 +3	1 2 3 4 5 6 7

Appendix 4.14
Participant information sheet, consent form, demographic questionnaire

Information Sheet for Participants

PLEASE READ THE FOLLOWING CAREFULLY

Study title: **An examination of the psychological characteristics and thought processes of competitive athletes**

What is the study about?
The purpose of this study is to examine the psychological characteristics and thought processes of competitive athletes.

At the University of Chichester before any project starts it has to be checked by university staff. They make sure that the research is good enough to carry out. This project has been checked by the staff of the Department of Sport and Exercise Science and has been passed by the University Ethics Committee.

Do you have to take part?
No! It is your choice whether you participate in the project or not. If at any time during the project you feel that you don't want to continue then you can tell the researcher that you want to stop – you do not have to give a reason.

What will you do in the project?
The study involves completing a series of psychological questionnaires. You will be asked to complete two questionnaires relating to your general personality in the first sitting and two questionnaires relating to your thoughts and feelings prior to competition in this second sitting. Please take your time to complete the questionnaires and answer as truthfully as possible.

In total, the first batch of questionnaires should take no longer than 8 minutes and the second batch of questionnaires should take no longer than 6 minutes so the total amount of time you will need to commit to this project is 14 minutes.

Why have you been asked to take part?

You have been asked to take part in this project since we need athletes who are competing in sport to gain an insight into the psychology of sport.

What happens to the information in the project?
All the information which is collected about you during the project will be kept private – the data will be identifiable for data analysis purposes but will only be observed by the primary researcher. Only group data will be used for publication purposes which will make it impossible to trace this information back to you individually.

Thank you for your time - please ask any questions if you are unsure or confused about what we have said.

What happens if you agree to take part?

- If you are happy to be involved in the project then please sign the informed consent form and the accompanied questionnaires.
- If you do not want to be involved in the project then thank you.

Who can you contact if you have any questions about the project?

Phil Birch – Email p.birch@chi.ac.uk: Phone 01243 816343 or
Dr Iain Greenlees - Email i.greenlees@chi.ac.uk: Phone 01243 816437

Thank you for your time

Consent Form

I understand that my participation in this project will involve completing two questionnaires which will assess my general personality and two questionnaires which will assess my thoughts and feelings prior to competing in sport. I understand that I will be asked to complete two questionnaires today and the remaining two questionnaires in one week's time approximately one hour prior competition. I understand that participation in this study is entirely voluntary and that I can withdraw from the study at any time without giving a reason.

I understand that I am free to ask any questions at any time. I am free to withdraw or discuss my concerns with Phil Birch (p.birch@chi.ac.uk) or Dr. Iain Greenlees (i.greenlees@chi.ac.uk).

I understand that the information I provide over the two collection periods will be identifiable for data analysis purposes **but will only be observed by the primary researcher**. I understand that only group data will be used for publication purposes which will make it impossible to trace this information back to me individually. I understand that this information will be stored confidentially and may be retained indefinitely.

I also understand that at the end of the study I will be provided with additional information and feedback about the purpose of the study.

I, ____(NAME) consent to participate in the study conducted by Phil Birch

Signed:

Date:

Demographic Questionnaire

Please fill in your details in the spaces provided

Age ______________

(years)

Gender Male ☐ Female ☐

Ethnic origin – Please indicate the most appropriate description of your ethnic origin.

(Please tick)

White - British	☐	Asian - Indian	☐
White - Irish	☐	Asian - Pakistani	☐
White – Any other White background	☐	Asian – Bangladeshi	☐
Mixed – White & Black Caribbean	☐	Black or Black British - Caribbean	☐
Mixed – White & Black African	☐	Black or Black British - African	☐
Mixed – White & Black Asian	☐	Black or Black British - Any other Black background	☐
Mixed – Any other Mixed background	☐	Chinese – Any Chinese background	☐
Any other ethnic background	☐		

Year of study (e.g., 1, 2, 3) ____

What sport are you going to be competing in? ____

How many years competitive playing experience do you have in this sport? ____________________________________

What is the highest level that you have played in this sport? ____

Appendix 4.15
Written debrief for the Revised Hardiness Confidence Questionnaire

Participant Debrief

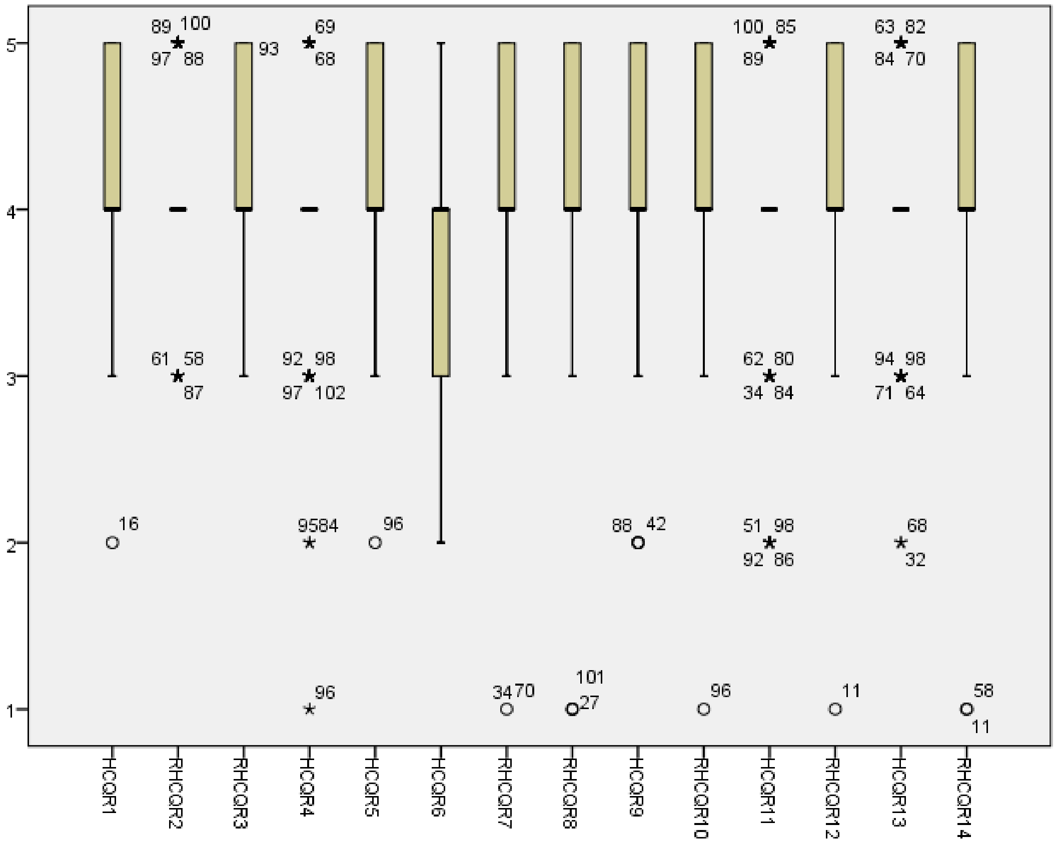
The aim of this study was to examine the suitability of the Revised Hardiness Confidence Questionnaire (HCQ-R) in measuring mental toughness in sport and to devise programmes to enhance the suitability of the HCQ-R.

This required us to compare the HCQ-R with a number of other psychological questionnaires (i.e., Social Desirability Scale, Immediate Anxiety Measures Scale, challenge and threat construal measure).

If you have any queries about this study please do not hesitate to contact Phil Birch (p.birch@chi.ac.uk) or Dr. Iain Greenlees (i.greenlees@chi.ac.uk).

Thank you again for your participation

Appendix 4.16
Box-plots for HCQ-R scores

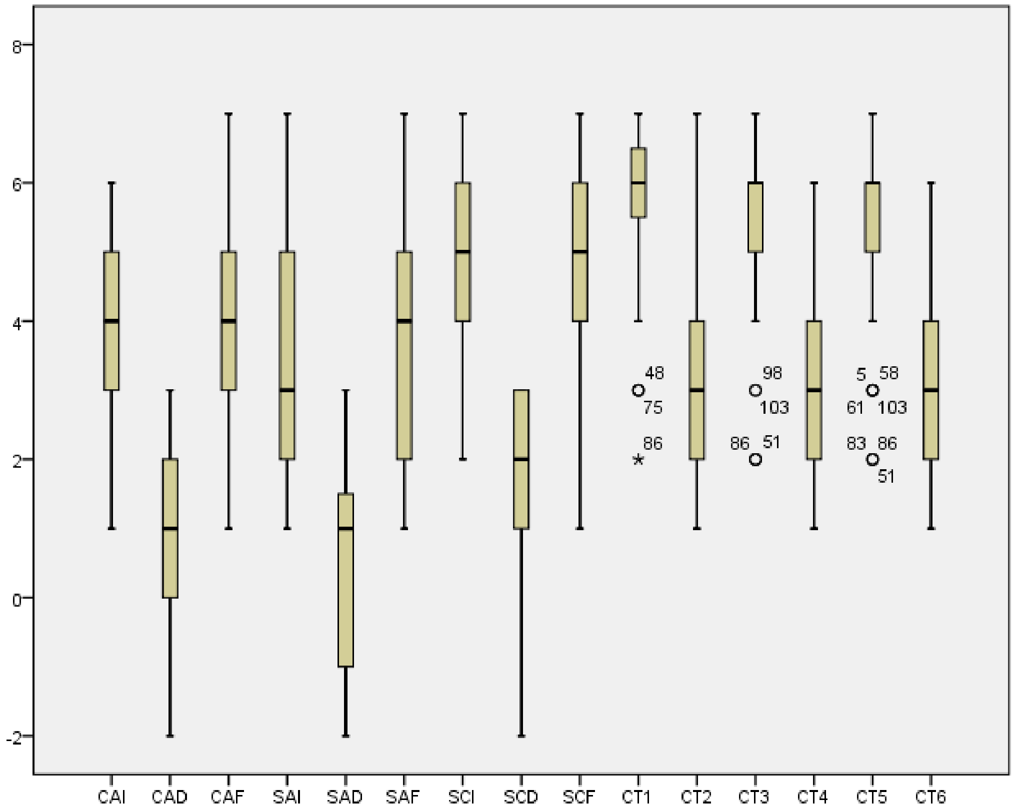


Appendix 4.17
Univariate normality statistics of the challenge and threat construal measure and the Immediate Anxiety Measurement Scale

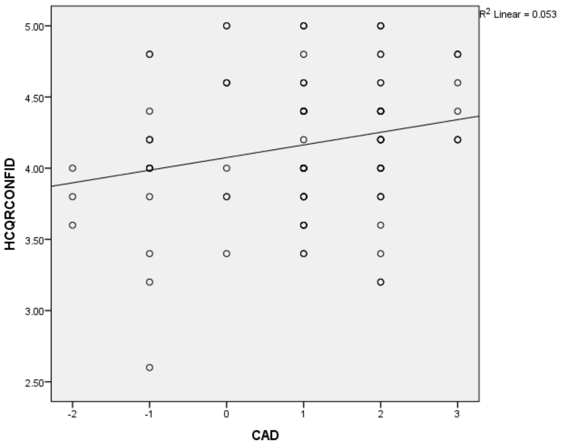
Univariate normality statistics of the challenge and threat construal measure and the IAMS

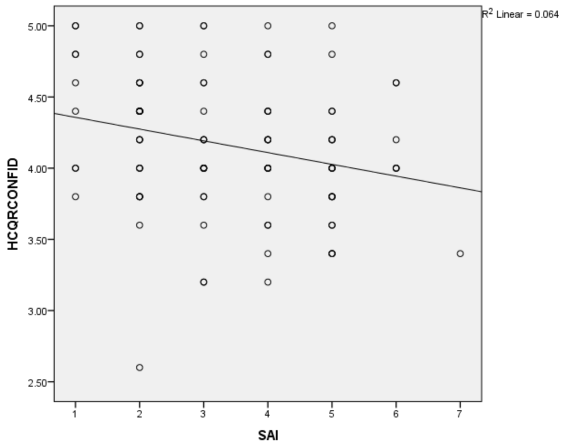
	Skewness ratio	Kurtosis ratio
CAI	-1.37	-1.36
CAD	-2.46	-0.66
CAF	0.10	-1.00
SAI	0.95	-1.87
SAD	-1.10	-1.80
SAF	0.13	-2.19
SCI	-2.37	-0.02
SCD	-5.04	1.25
SCF	-3.41	1.47
CT1	-5.60	5.46
CT2	2.37	-0.56
CT3	-5.50	4.51
CT4	1.98	-1.28
CT5	-5.58	4.05
CT6	2.10	-0.64
PRECHAL	-6.85	7.66
PRETHREAT	1.62	-0.28

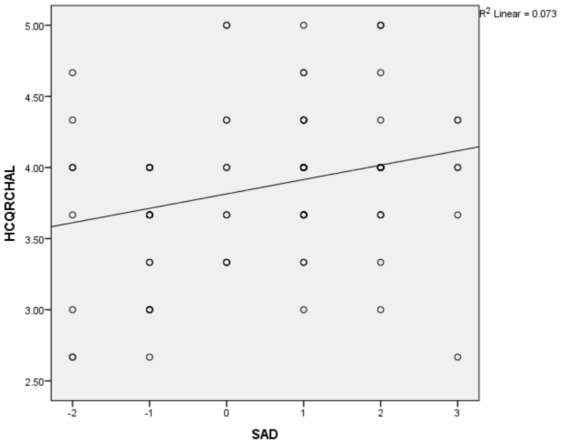
Appendix 4.18
Box-plots for the challenge and threat construal measure and IAMS scores

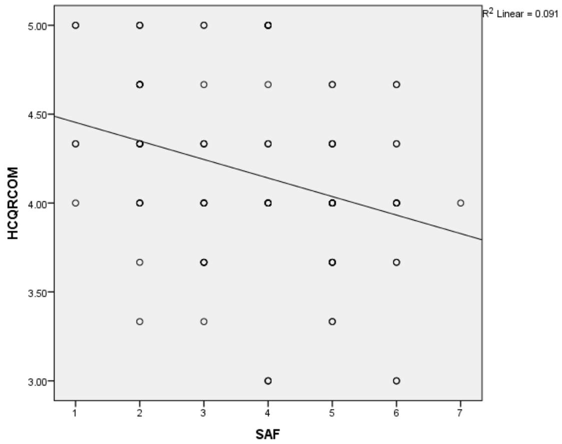


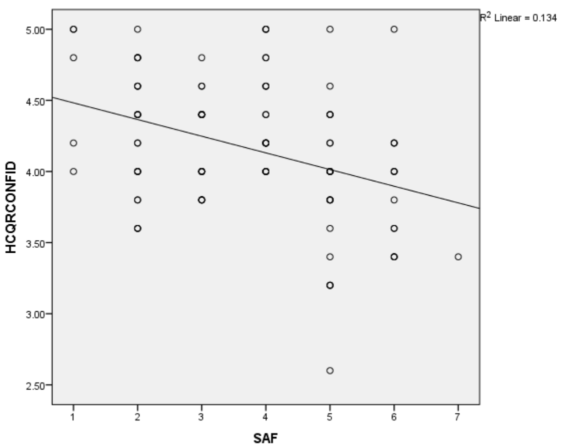
Appendix 4.19
Scatterplots between respective subscales of the HCQ-R and the challenge and threat construal measure and the IAMS

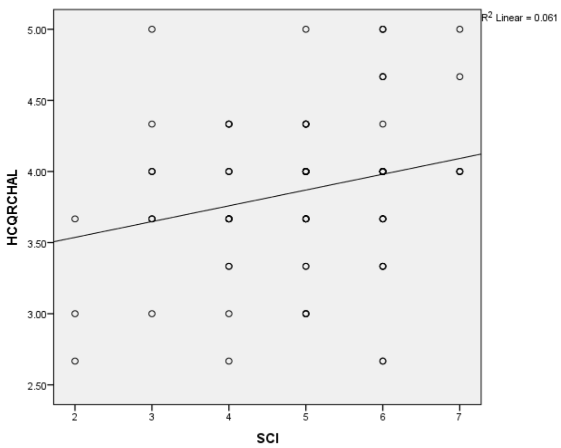


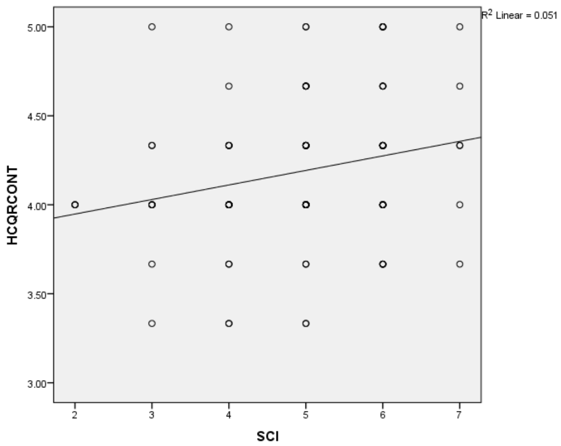


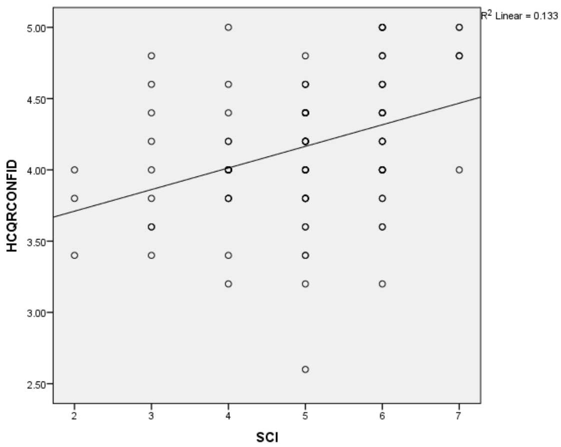


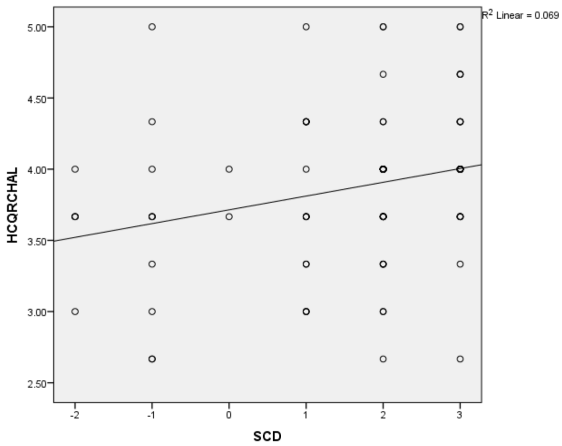


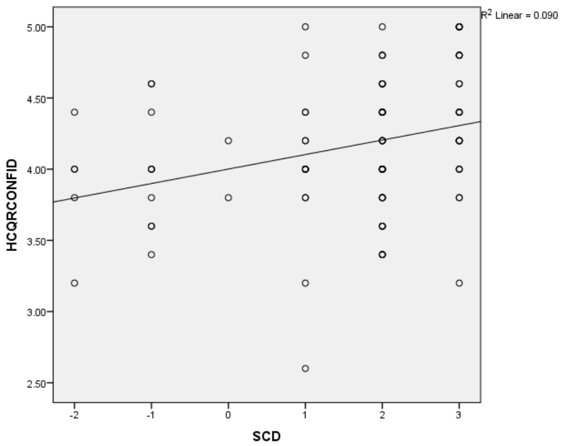


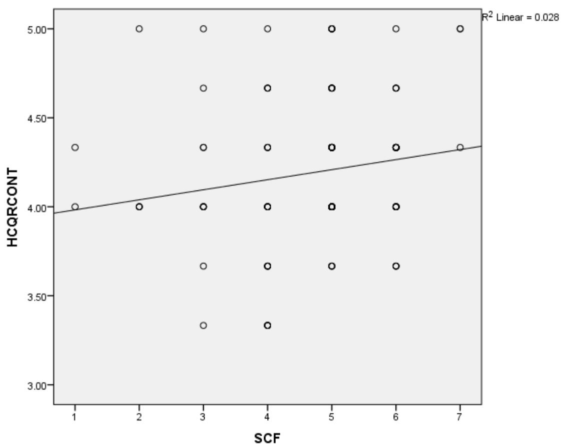


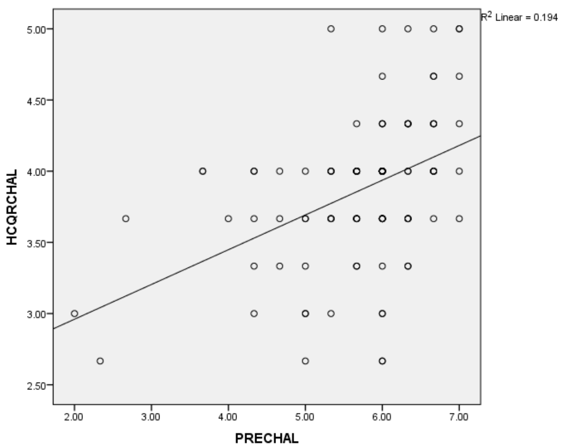

R² Linear = 0.051
HCQRCONT
SCI
5.00
4.50
4.00
3.50
3.00
2
3
4
5
6
7

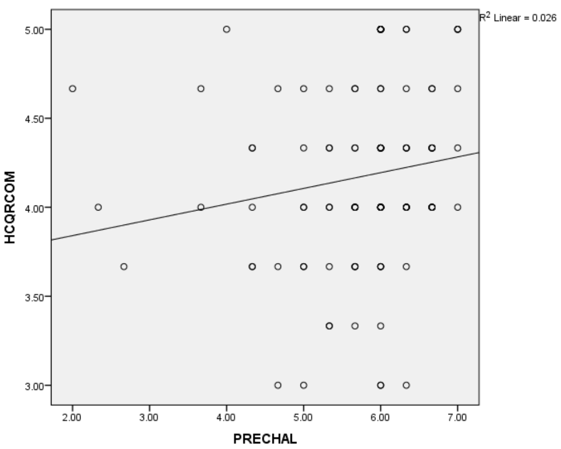

R² Linear = 0.133
HCQRCONFID
SCI
5.00
4.50
4.00
3.50
3.00
2.50
2
3
4
5
6
7

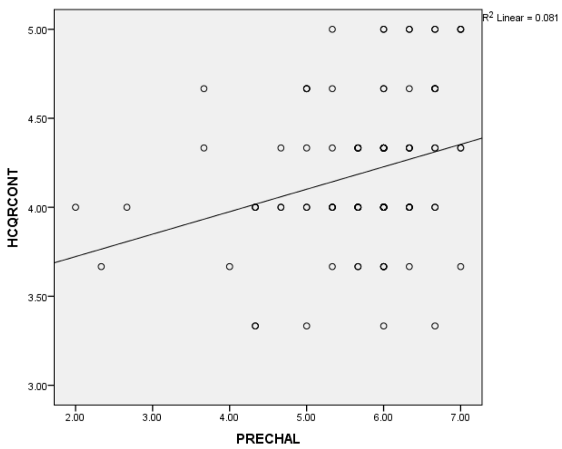

R² Linear = 0.069
HCQRCHAL
SCD
5.00
4.50
4.00
3.50
3.00
2.50
-2
-1
0
1
2
3

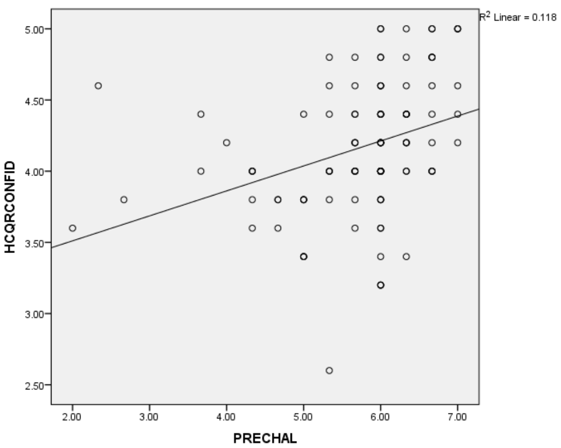

R² Linear = 0.090
HCQRCONFID
SCD
5.00
4.50
4.00
3.50
3.00
2.50
-2
-1
0
1
2
3

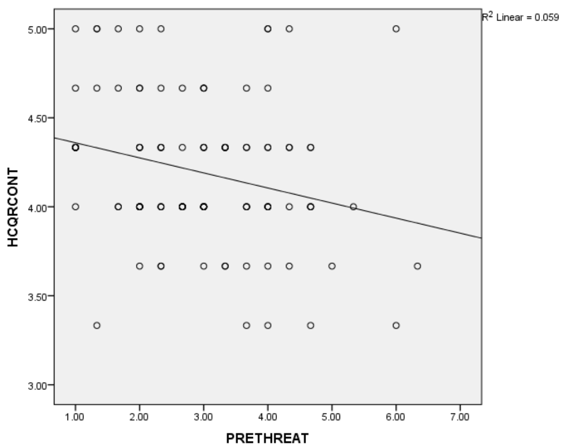

R² Linear = 0.028
HCQRCONT
SCF
5.00
4.50
4.00
3.50
3.00
1
2
3
4
5
6
7

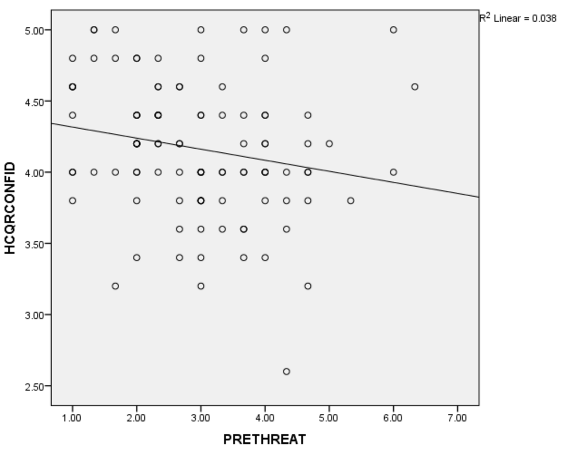

R² Linear = 0.194
HCQRCHAL
PRECHAL
5.00
4.50
4.00
3.50
3.00
2.50
2.00
3.00
4.00
5.00
6.00
7.00


R2 Linear = 0.026
HCQRCOM
PRECHAL


R2 Linear = 0.081
HCQRCONT
PRECHAL


R2 Linear = 0.118
HCQRCONFID
PRECHAL


R2 Linear = 0.059
HCQRCONT
PRETHREAT


R2 Linear = 0.038
HCQRCONFID
PRETHREAT

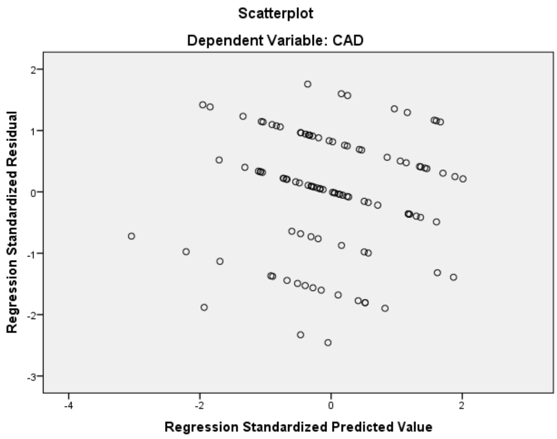
Appendix 4.20
Mahalanobis distance statistics for HCQ-R scores
(CAD regression model exemplar)

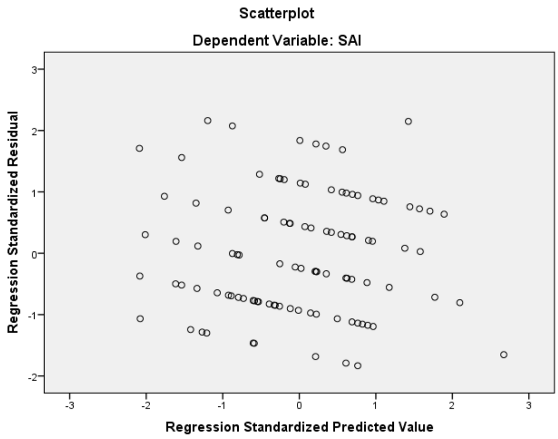
Residuals Statistics[a]

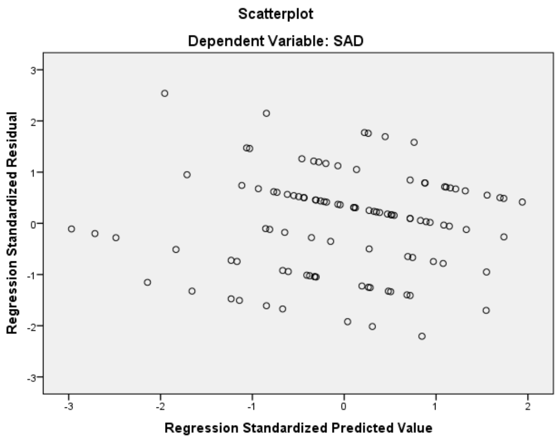
	Minimum	Maximum	Mean	Std. Deviation	N
Predicted Value	-.12	1.74	1.00	.370	104
Std. Predicted Value	-3.043	2.010	.000	1.000	104
Standard Error of Predicted Value	.137	.463	.255	.076	104
Adjusted Predicted Value	.02	1.78	1.00	.373	104
Residual	-2.983	2.132	.000	1.190	104
Std. Residual	-2.457	1.756	.000	.980	104
Stud. Residual	-2.475	1.790	-.001	1.004	104
Deleted Residual	-3.027	2.216	-.003	1.249	104
Stud. Deleted Residual	-2.542	1.811	-.004	1.012	104
Mahal. Distance	.327	**13.962**	3.962	2.881	104
Cook's Distance	.000	.098	.010	.015	104
Centered Leverage Value	.003	.136	.038	.028	104

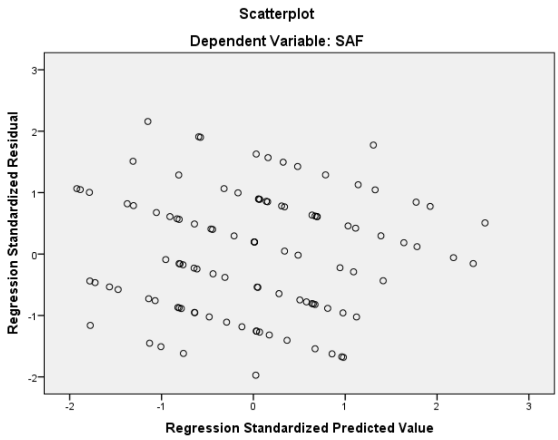
a. Dependent Variable: CAD

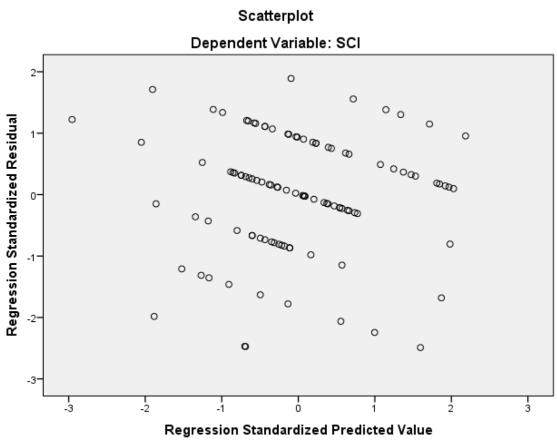
Appendix 4.21
Scatterplots between the regression standardised residual and the regression standardised predicted value of each dependent variable tested

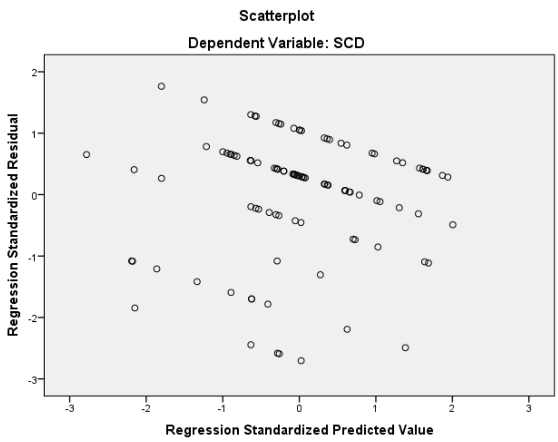


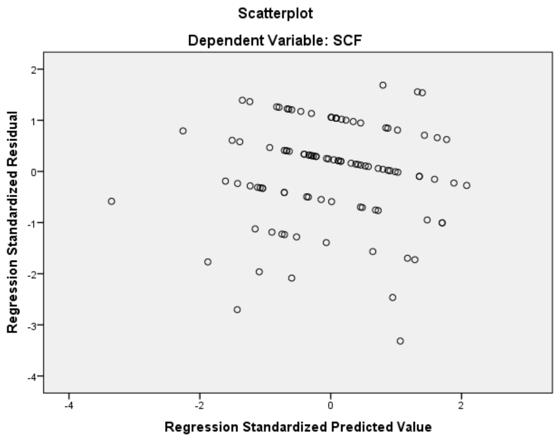


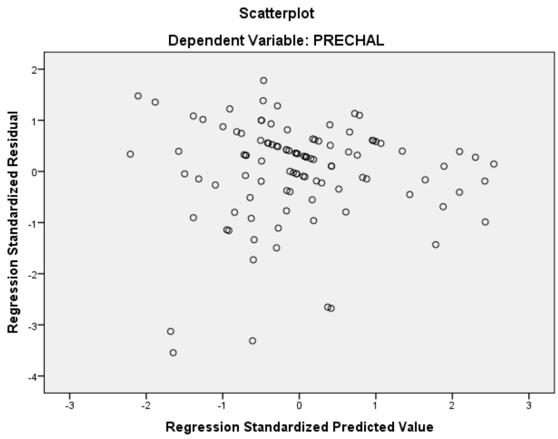


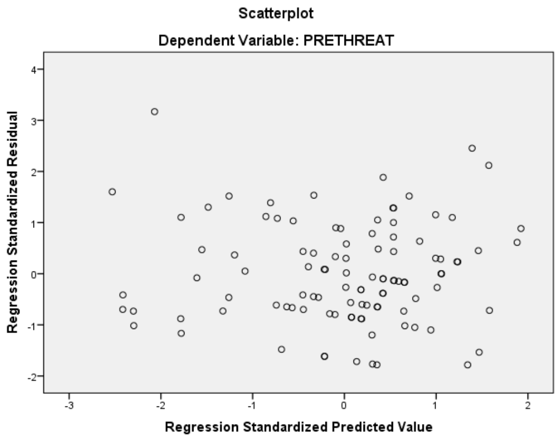







Scatterplot
Dependent Variable: SCF
Regression Standardized Residual
Regression Standardized Predicted Value


Scatterplot
Dependent Variable: PRECHAL
Regression Standardized Residual
Regression Standardized Predicted Value


Scatterplot
Dependent Variable: PRETHREAT
Regression Standardized Residual
Regression Standardized Predicted Value

Appendix 4.22
Multiple hierarchical regression models

CAD

Variables Entered/Removed[a]

Model	Variables Entered	Variables Removed	Method
1	HCQRCONT[b]	.	Enter
2	HCQRCHAL, HCQRCONFID[b]	.	Enter
3	HCQRCOM[b]	.	Enter

a. Dependent Variable: CAD

b. All requested variables entered.

Model Summary[d]

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.187[a]	.035	.025	1.230
2	.258[b]	.066	.038	1.222
3	.296[c]	.088	.051	1.214

a. Predictors: (Constant), HCQRCONT

b. Predictors: (Constant), HCQRCONT, HCQRCHAL, HCQRCONFID

c. Predictors: (Constant), HCQRCONT, HCQRCHAL, HCQRCONFID, HCQRCOM

d. Dependent Variable: CAD

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	5.577	1	5.577	3.684	.058[b]
	Residual	154.423	102	1.514		
	Total	160.000	103			
2	Regression	10.634	3	3.545	2.373	.075[c]
	Residual	149.366	100	1.494		
	Total	160.000	103			
3	Regression	14.065	4	3.516	2.385	.056[d]
	Residual	145.935	99	1.474		
	Total	160.000	103			

a. Dependent Variable: CAD

b. Predictors: (Constant), HCQRCONT

c. Predictors: (Constant), HCQRCONT, HCQRCHAL, HCQRCONFID

d. Predictors: (Constant), HCQRCONT, HCQRCHAL, HCQRCONFID, HCQRCOM

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	Collinearity Statistics	
		B	Std. Error	Beta			Tolerance	VIF
1	(Constant)	-1.338	1.224		-1.093	.277		
	HCQRCONT	.558	.291	.187	1.919	.058	1.000	1.000
2	(Constant)	-2.622	1.415		-1.853	.067		
	HCQRCONT	.326	.316	.109	1.033	.304	.837	1.195
	HCQRCHAL	.111	.262	.046	.424	.672	.779	1.284
	HCQRCONFID	.438	.294	.169	1.492	.139	.731	1.367
3	(Constant)	-1.705	1.528		-1.116	.267		
	HCQRCONT	.312	.314	.104	.995	.322	.836	1.196
	HCQRCHAL	.178	.264	.074	.675	.501	.757	1.321
	HCQRCONFID	.552	.301	.212	1.831	.070	.687	1.456
	HCQRCOM	-.381	.250	-.158	-1.526	.130	.857	1.166

a. Dependent Variable: CAD

SAI

Variables Entered/Removed[a]

Model	Variables Entered	Variables Removed	Method
1	HCQRCONFID[b]	.	Enter
2	HCQRCOM, HCQRCONT, HCQRCHAL[b]	.	Enter

a. Dependent Variable: SAI

b. All requested variables entered.

Model Summary[c]

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.253[a]	.064	.055	1.426
2	.265[b]	.070	.033	1.442

a. Predictors: (Constant), HCQRCONFID

b. Predictors: (Constant), HCQRCONFID, HCQRCOM, HCQRCONT, HCQRCHAL

c. Dependent Variable: SAI

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	14.130	1	14.130	6.949	.010[b]
	Residual	207.409	102	2.033		
	Total	221.538	103			
2	Regression	15.552	4	3.888	1.869	.122[c]
	Residual	205.986	99	2.081		
	Total	221.538	103			

a. Dependent Variable: SAI

b. Predictors: (Constant), HCQRCONFID

c. Predictors: (Constant), HCQRCONFID, HCQRCOM, HCQRCONT, HCQRCHAL

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	Collinearity Statistics	
		B	Std. Error	Beta			Tolerance	VIF
1	(Constant)	6.564	1.229		5.343	.000		
	HCQRCONFID	-.773	.293	-.253	-2.636	.010	1.000	1.000
2	(Constant)	7.400	1.816		4.075	.000		
	HCQRCONFID	-.662	.358	-.216	-1.849	.067	.687	1.456
	HCQRCOM	-.234	.297	-.083	-.789	.432	.857	1.166
	HCQRCHAL	.006	.314	.002	.018	.986	.757	1.321
	HCQRCONT	-.082	.373	-.023	-.221	.826	.836	1.196

a. Dependent Variable: SAI

SAD

Variables Entered/Removed[a]

Model	Variables Entered	Variables Removed	Method
1	HCQRCONT[b]	.	Enter
2	HCQRCHAL, HCQRCONFID[b]	.	Enter
3	HCQRCOM[b]	.	Enter

a. Dependent Variable: SAD

b. All requested variables entered.

Model Summary[d]

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.052[a]	.003	-.007	1.391
2	.315[b]	.099	.072	1.335
3	.338[c]	.114	.079	1.330

a. Predictors: (Constant), HCQRCONT

b. Predictors: (Constant), HCQRCONT, HCQRCHAL, HCQRCONFID

c. Predictors: (Constant), HCQRCONT, HCQRCHAL, HCQRCONFID, HCQRCOM

d. Dependent Variable: SAD

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	.537	1	.537	.278	.599[b]
	Residual	197.309	102	1.934		
	Total	197.846	103			
2	Regression	19.638	3	6.546	3.673	.015[c]
	Residual	178.208	100	1.782		
	Total	197.846	103			
3	Regression	22.627	4	5.657	3.196	.016[d]
	Residual	175.219	99	1.770		
	Total	197.846	103			

a. Dependent Variable: SAD

b. Predictors: (Constant), HCQRCONT

c. Predictors: (Constant), HCQRCONT, HCQRCHAL, HCQRCONFID

d. Predictors: (Constant), HCQRCONT, HCQRCHAL, HCQRCONFID, HCQRCOM

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	Collinearity Statistics	
		B	Std. Error	Beta			Tolerance	VIF
1	(Constant)	1.264	1.384		.913	.363		
	HCQRCONT	-.173	.328	-.052	-.527	.599	1.000	1.000
2	(Constant)	-.217	1.545		-.140	.889		
	HCQRCONT	-.391	.345	-.118	-1.135	.259	.837	1.195
	HCQRCHAL	.932	.287	.350	3.253	.002	.779	1.284
	HCQRCONFID	-.291	.321	-.101	-.907	.366	.731	1.367
3	(Constant)	.639	1.675		.382	.704		
	HCQRCONT	-.404	.344	-.122	-1.176	.242	.836	1.196
	HCQRCHAL	.995	.290	.373	3.435	.001	.757	1.321
	HCQRCONFID	-.186	.330	-.064	-.562	.575	.687	1.456
	HCQRCOM	-.356	.274	-.133	-1.300	.197	.857	1.166

a. Dependent Variable: SAD

Variables Entered/Removed[a]

Model	Variables Entered	Variables Removed	Method
1	HCQRCONFID[b]	.	Enter
2	HCQRCOM, HCQRCONT, HCQRCHAL[b]	.	Enter

a. Dependent Variable: SAF

b. All requested variables entered.

Model Summary[c]

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.366[a]	.134	.126	1.400
2	.417[b]	.174	.141	1.388

a. Predictors: (Constant), HCQRCONFID

b. Predictors: (Constant), HCQRCONFID, HCQRCOM, HCQRCONT, HCQRCHAL

c. Dependent Variable: SAF

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	30.946	1	30.946	15.785	.000[b]
	Residual	199.967	102	1.960		
	Total	230.913	103			
2	Regression	40.160	4	10.040	5.211	.001[c]
	Residual	190.754	99	1.927		
	Total	230.913	103			

a. Dependent Variable: SAF

b. Predictors: (Constant), HCQRCONFID

c. Predictors: (Constant), HCQRCONFID, HCQRCOM, HCQRCONT, HCQRCHAL

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	Collinearity Statistics	
		B	Std. Error	Beta			Tolerance	VIF
1	(Constant)	8.483	1.206		7.032	.000		
	HCQRCONFID	-1.144	.288	-.366	-3.973	.000	1.000	1.000
2	(Constant)	9.232	1.747		5.283	.000		
	HCQRCONFID	-1.021	.344	-.327	-2.966	.004	.687	1.456
	HCQRCHAL	.009	.302	.003	.030	.976	.757	1.321
	HCQRCOM	-.578	.286	-.200	-2.025	.046	.857	1.166
	HCQRCONT	.267	.358	.074	.744	.459	.836	1.196

a. Dependent Variable: SAF

Variables Entered/Removed[a]

Model	Variables Entered	Variables Removed	Method
1	HCQRCONFID[b]	.	Enter
2	HCQRCOM, HCQRCONT, HCQRCHAL[b]	.	Enter

a. Dependent Variable: SCI

b. All requested variables entered.

Model Summary[c]

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.365[a]	.133	.125	1.079
2	.386[b]	.149	.115	1.085

a. Predictors: (Constant), HCQRCONFID

b. Predictors: (Constant), HCQRCONFID, HCQRCOM, HCQRCONT, HCQRCHAL

c. Dependent Variable: SCI

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	18.221	1	18.221	15.648	.000[b]
	Residual	118.770	102	1.164		
	Total	136.990	103			
2	Regression	20.444	4	5.111	4.341	.003[c]
	Residual	116.547	99	1.177		
	Total	136.990	103			

a. Dependent Variable: SCI

b. Predictors: (Constant), HCQRCONFID

c. Predictors: (Constant), HCQRCONFID, HCQRCOM, HCQRCONT, HCQRCHAL

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	Collinearity Statistics	
		B	Std. Error	Beta			Tolerance	VIF
1	(Constant)	1.337	.930		1.438	.154		
	HCQRCONFID	.878	.222	.365	3.956	.000	1.000	1.000
2	(Constant)	.477	1.366		.349	.728		
	HCQRCONFID	.724	.269	.301	2.692	.008	.687	1.456
	HCQRCHAL	.210	.236	.095	.891	.375	.757	1.321
	HCQRCOM	-.082	.223	-.037	-.366	.715	.857	1.166
	HCQRCONT	.244	.280	.088	.872	.386	.836	1.196

a. Dependent Variable: SCI

Variables Entered/Removed[a]

Model	Variables Entered	Variables Removed	Method
1	HCQRCONFID[b]	.	Enter
2	HCQRCOM, HCQRCONT, HCQRCHAL[b]	.	Enter

a. Dependent Variable: SCD

b. All requested variables entered.

Model Summary[c]

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.300[a]	.090	.081	1.353
2	.374[b]	.140	.105	1.335

a. Predictors: (Constant), HCQRCONFID

b. Predictors: (Constant), HCQRCONFID, HCQRCOM, HCQRCONT, HCQRCHAL

c. Dependent Variable: SCD

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	18.400	1	18.400	10.056	.002[b]
	Residual	186.639	102	1.830		
	Total	205.038	103			
2	Regression	28.642	4	7.161	4.019	.005[c]
	Residual	176.396	99	1.782		
	Total	205.038	103			

a. Dependent Variable: SCD

b. Predictors: (Constant), HCQRCONFID

c. Predictors: (Constant), HCQRCONFID, HCQRCOM, HCQRCONT, HCQRCHAL

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	Collinearity Statistics	
		B	Std. Error	Beta			Tolerance	VIF
1	(Constant)	-2.075	1.165		-1.781	.078		
	HCQRCONFID	.882	.278	.300	3.171	.002	1.000	1.000
2	(Constant)	-1.886	1.680		-1.123	.264		
	HCQRCONFID	.796	.331	.270	2.404	.018	.687	1.456
	HCQRCHAL	.510	.291	.188	1.755	.082	.757	1.321
	HCQRCOM	-.503	.275	-.184	-1.830	.070	.857	1.166
	HCQRCONT	.069	.345	.020	.201	.841	.836	1.196

a. Dependent Variable: SCD

SCF

Variables Entered/Removed[a]

Model	Variables Entered	Variables Removed	Method
1	HCQRCONFID[b]	.	Enter
2	HCQRCOM, HCQRCONT, HCQRCHAL[b]	.	Enter

a. Dependent Variable: SCF

b. All requested variables entered.

Model Summary[c]

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.190[a]	.036	.027	1.212
2	.244[b]	.059	.021	1.215

a. Predictors: (Constant), HCQRCONFID

b. Predictors: (Constant), HCQRCONFID, HCQRCOM, HCQRCONT, HCQRCHAL

c. Dependent Variable: SCF

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	5.593	1	5.593	3.810	.054[b]
	Residual	149.753	102	1.468		
	Total	155.346	103			
2	Regression	9.213	4	2.303	1.560	.191[c]
	Residual	146.134	99	1.476		
	Total	155.346	103			

a. Dependent Variable: SCF

b. Predictors: (Constant), HCQRCONFID

c. Predictors: (Constant), HCQRCONFID, HCQRCOM, HCQRCONT, HCQRCHAL

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	Collinearity Statistics	
		B	Std. Error	Beta			Tolerance	VIF
1	(Constant)	2.687	1.044		2.574	.011		
	HCQRCONFID	.486	.249	.190	1.952	.054	1.000	1.000
2	(Constant)	2.573	1.530		1.683	.096		
	HCQRCONFID	.509	.301	.198	1.688	.095	.687	1.456
	HCQRCHAL	-.067	.265	-.028	-.252	.801	.757	1.321
	HCQRCOM	-.270	.250	-.114	-1.080	.283	.857	1.166
	HCQRCONT	.335	.314	.114	1.068	.288	.836	1.196

a. Dependent Variable: SCF

CHALLENGE

Variables Entered/Removed[a]

Model	Variables Entered	Variables Removed	Method
1	HCQRCHAL[b]	.	Enter
2	HCQRCONT, HCQRCONFID[b]	.	Enter
3	HCQRCOM[b]	.	Enter

a. Dependent Variable: PRECHAL

b. All requested variables entered.

Model Summary[d]

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.441[a]	.194	.186	.84742
2	.485[b]	.235	.212	.83404
3	.485[c]	.235	.204	.83823

a. Predictors: (Constant), HCQRCHAL

b. Predictors: (Constant), HCQRCHAL, HCQRCONT, HCQRCONFID

c. Predictors: (Constant), HCQRCHAL, HCQRCONT, HCQRCONFID, HCQRCOM

d. Dependent Variable: PRECHAL

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	17.666	1	17.666	24.600	.000[b]
	Residual	73.248	102	.718		
	Total	90.913	103			
2	Regression	21.351	3	7.117	10.231	.000[c]
	Residual	69.562	100	.696		
	Total	90.913	103			
3	Regression	21.352	4	5.338	7.597	.000[d]
	Residual	69.561	99	.703		
	Total	90.913	103			

a. Dependent Variable: PRECHAL

b. Predictors: (Constant), HCQRCHAL

c. Predictors: (Constant), HCQRCHAL, HCQRCONT, HCQRCONFID

d. Predictors: (Constant), HCQRCHAL, HCQRCONT, HCQRCONFID, HCQRCOM

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	Collinearity Statistics	
		B	Std. Error	Beta			Tolerance	VIF
1	(Constant)	2.641	.627		4.216	.000		
	HCQRCHAL	.796	.161	.441	4.960	.000	1.000	1.000
2	(Constant)	.972	.965		1.006	.317		
	HCQRCHAL	.611	.179	.338	3.412	.001	.779	1.284
	HCQRCONFID	.279	.201	.142	1.390	.168	.731	1.367
	HCQRCONT	.292	.215	.130	1.358	.177	.837	1.195
3	(Constant)	.990	1.055		.938	.350		
	HCQRCHAL	.612	.182	.339	3.355	.001	.757	1.321
	HCQRCONFID	.281	.208	.143	1.351	.180	.687	1.456
	HCQRCONT	.292	.216	.130	1.349	.180	.836	1.196
	HCQRCOM	-.008	.172	-.004	-.045	.964	.857	1.166

a. Dependent Variable: PRECHAL

THREAT

Variables Entered/Removed[a]

Model	Variables Entered	Variables Removed	Method
1	HCQRCHAL[b]	.	Enter
2	HCQRCONT, HCQRCONFID[b]	.	Enter
3	HCQRCOM[b]	.	Enter

a. Dependent Variable: PRETHREAT

b. All requested variables entered.

Model Summary[d]

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.192[a]	.037	.027	1.18548
2	.282[b]	.080	.052	1.17026
3	.282[c]	.080	.043	1.17615

a. Predictors: (Constant), HCQRCHAL

b. Predictors: (Constant), HCQRCHAL, HCQRCONT, HCQRCONFID

c. Predictors: (Constant), HCQRCHAL, HCQRCONT, HCQRCONFID, HCQRCOM

d. Dependent Variable: PRETHREAT

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	5.473	1	5.473	3.894	.051[b]
	Residual	143.348	102	1.405		
	Total	148.821	103			
2	Regression	11.871	3	3.957	2.889	.039[c]
	Residual	136.950	100	1.370		
	Total	148.821	103			
3	Regression	11.871	4	2.968	2.145	.081[d]
	Residual	136.950	99	1.383		
	Total	148.821	103			

a. Dependent Variable: PRETHREAT

b. Predictors: (Constant), HCQRCHAL

c. Predictors: (Constant), HCQRCHAL, HCQRCONT, HCQRCONFID

d. Predictors: (Constant), HCQRCHAL, HCQRCONT, HCQRCONFID, HCQRCOM

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	Collinearity Statistics	
		B	Std. Error	Beta			Tolerance	VIF
1	(Constant)	4.689	.876		5.350	.000		
	HCQRCHAL	-.443	.225	-.192	-1.973	.051	1.000	1.000
2	(Constant)	6.931	1.355		5.117	.000		
	HCQRCHAL	-.234	.251	-.101	-.933	.353	.779	1.284
	HCQRCONT	-.529	.302	-.184	-1.751	.083	.837	1.195
	HCQRCONFID	-.200	.281	-.080	-.710	.479	.731	1.367
3	(Constant)	6.926	1.481		4.678	.000		
	HCQRCHAL	-.235	.256	-.102	-.917	.362	.757	1.321
	HCQRCONT	-.529	.304	-.184	-1.741	.085	.836	1.196
	HCQRCONFID	-.200	.292	-.080	-.687	.494	.687	1.456
	HCQRCOM	.002	.242	.001	.008	.994	.857	1.166

a. Dependent Variable: PRETHREAT