The Concept of Mental Toughness:
Tests of Dimensionality, Nomological Network and Traitness

Daniel F. Gucciardi\textsuperscript{1, 2}, Sheldon Hanton\textsuperscript{3}, Sandy Gordon\textsuperscript{4}, Clifford J. Mallett\textsuperscript{2} and Philip Temby\textsuperscript{5}

\textsuperscript{1}School of Physiotherapy and Exercise Science, Curtin University
\textsuperscript{2}School of Human Movement Studies, The University of Queensland
\textsuperscript{3}Cardiff School of Sport, Cardiff Metropolitan University
\textsuperscript{4}School of Sport Science, Exercise, and Health, The University of Western Australia
\textsuperscript{5}Land Operations Division, Defence Science and Technology Organization, South Australia

Authors’ Notes: This research was part funded by a University of Queensland Early Career Researcher Grant and Postdoctoral Research Fellowship to Daniel Gucciardi.

Appreciation is extended to Jon Chan for his assistance with data collection for Study 2, and to all those people and organizations that assisted with participant recruitment.

Corresponding Author: Daniel Gucciardi, School of Physiotherapy and Exercise Science, Curtin University, GPO Box U1987, Perth, Australia, 6845. Phone: +61 8 9266 3653, Fax: +61 8 9266 3699, Email: daniel.f.gucciardi@gmail.com

This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the Version of Record. Please cite this article as doi: 10.1111/jopy.12079

This article is protected by copyright. All rights reserved.
Abstract

Objectives: Mental toughness has received increased scholarly attention in recent years, yet conceptual issues related to its (i) dimensionality, (ii) nomological network, and (iii) traitness remain unresolved. The series of studies reported in this paper were designed to examine these three substantive issues across several achievement contexts including sport, education, military and the workplace. Methods: Five studies were conducted to examine these research aims: Study 1: N = 30; Study 2: calibration sample (n = 418), tertiary students (n = 500), athletes (n = 427), and employees (n = 550); Study 3: N = 497 employees; Study 4: N = 203 tertiary students; Study 5: N = 115 army candidates. Results: Collectively, the results of these studies revealed that mental toughness may be best conceptualized as a undimensional rather than a multidimensional concept; plays an important role in performance, goal progress, and thriving despite stress; and can vary and have enduring properties across situations and time. Conclusion: This series of studies provides a foundation for further basic and applied research of mental toughness across various achievement contexts.

Keywords: diary study; goal striving; multilevel structural equation modeling; personal resources; resilience; stress
Facing adversity and dealing with challenges is commonplace in life’s stressful achievement contexts. Students, for example, have to deal with the increased self-responsibility associated with their transition into tertiary education as well as perform under the pressure of a final exam. Businesspeople are under constant pressure to achieve key performance indicators to ensure that they remain employed in an unstable economic climate. Athletes often have to perform to their best under some of the most physically (e.g., fatigue) and mentally (e.g., crowd pressure) demanding circumstances. Across most achievement settings mental toughness (MT) is commonly referred to as the defining attribute that enables one to thrive in demanding situations (Jones & Moorhouse, 2007; Weinberg, 2010). Thus, it is not surprising that MT is a popular and highly valued concept, particularly within those contexts (e.g., business, education, medicine, military) where high performance underpins innovation, success, and competitive advantage (Asken, Grossman, & Christensen, 2010; Jones & Moorhouse, 2007). Some authors have gone as far to suggest that MT is fundamental to success in life (Wakefield, 2008; Weinberg, 2010).

The large volume of books within the popular press and news both in print and online suggests that MT is a concept that is meaningful for achievement and, more broadly, individuals’ lives. The popularity of MT among the general public has inevitably led to the upsurge in academic attention that has occurred over the past decade. For example, an October 3rd 2013 search of the Web of Knowledge database from 1900 to 2013 revealed 145 papers, chapters, or conference presentations in which the term MT appears in the title or topic; over 95% of which have surfaced since 2000. The vast majority of this research has been conducted within sport contexts, relying initially on retrospective interviews with elite athletes and coaches, and then cross-sectional interviews or surveys (for reviews, see Gucciardi & Gordon, 2011). In recent years, scholars have expanded their focus beyond the sporting field to achievement contexts such as surgery (Colbert, Scott, Dale, & Brennan,
Although the past decade has evidenced increased scholarly attention to MT, a number of conceptual and methodological concerns have limited the usefulness of these findings for conceptual development. First, the empirical focus on MT primarily within sport contexts limits the extent to which these findings generalize to other samples. Second, when MT has been examined in non-sport contexts, scholars have applied sport models with little explanation of the substantive or empirical evidence for doing so; there is evidence that it is sometimes erroneous to make this assumption (Gucciardi, Hanton, & Mallett, 2012). Third, with few exceptions (i.e., Hardy, Bell, & Beattie, in press), the area is characterized by piecemeal investigations in which the findings of individual studies have not been adequately integrated into a coherent conceptualization of MT. In particular, there have been at least 10 qualitative investigations designed to delineate the key features of mental toughness, yet there has been no attempt to systematically synthesize this research. This apparent disconnect between empirical research and concept development or refinement with MT is at odds with the aim of substantive parsimony. There is still a need for a clearer definition and conceptualization that underpins concept development and subsequent testing of the substantive propositions of MT. Accordingly, the overarching goal of this research was to revisit the fundamental question “what is this thing called MT” (cf. Jones, Hanton, & Connaughton, 2002) via a series of systematic and theoretically integrated investigations.

Conceptual Foundations

As with many research topics in their relatively early stages of development, conceptual clarity is a problem for the field of MT because there is yet no clear consensus as to what it really is and what it is not. Key scholarly definitions of MT derived from empirical research are detailed in Table 1. These definitions of MT have sparked independent streams...
of empirical research (for reviews, see Gucciardi & Gordon, 2011), yet more agreement on common elements of a concept is required before such bodies of knowledge can contribute to its evolution (Suddaby, 2010).

Owing to the lack of an established or empirically supported conceptual framework underpinning current conceptualizations of MT, we initially took an inductive approach to guide the development of our conceptual model (cf. Locke, 2007). Our conceptual model and associated working definition is based on a synthesis of the existing empirical evidence obtained from a comprehensive and up-to-date review of the literature (e.g., Gucciardi & Gordon, 2011), an electronic database search of articles that have been through a peer-review process (i.e., PsycINFO, Web of Science, Google Scholar), and unpublished interviews we have conducted with over 30 performers from non-sport contexts (i.e., military, surgery, business, tertiary education). In an effort to summarize and integrate these efforts, we propose a working definition of MT because it may be refined and extended over time as new findings about its nature emerge. Because MT is an umbrella term used to group a number of related concepts, it makes sense that a working definition would initially be relatively broad in nature, with the corresponding model providing specificity regarding the nature and number of dimensions (Bacharach, 1989). Thus, we define MT as a personal capacity to produce consistently high levels of subjective (e.g., personal goals or strivings) or objective performance (e.g., sales, race time, GPA) despite everyday challenges and stressors as well as significant adversities.

Our working conceptualization appears to adequately integrate key features of academic descriptions of MT with dictionary definitions of “mental” and “tough” detailed in Table 1 (cf. Locke, 2003). Central to both academic and English language definitions is the personal nature of MT – it is a quality that resides within an individual. Also inherent within these definitions is the notion of being able to withstand or endure challenging or adverse
situations. Although not explicitly captured in the English language definitions, the centrality of this personal quality for high performance and/or goal attainment is central to academic descriptions of MT. For example, Jones et al. (2002) linked MT with being “more consistent and better than your opponents” (p. 201), whereas Coulter et al. (2010) described it as being the foundation upon which to “consistently achieve [your] goals” (p. 715).

Unlike previous definitions that have encompassed examples of what is included in the construct (e.g., Clough et al., 2002) or are other-dependent such as “cope better than your opponents” (e.g., Jones et al., 2002, p. 209), our relatively broad description of MT includes several important assumptions that we believe provide a useful foundation upon which to help clarify its meaning. First, we contend that MT represents a “resource caravan” (Hobfoll, 2002, p. 312) or aggregation of several personal resources or assets which are common to people who share social and environmental conditions (e.g., biology, culture); that is, these personal resources or assets are interwoven, with the common conceptually identifiable element among them being a process driving consistently high levels of subjective or objective performance. Second, an inherent feature of this conceptualization is that MT resides within an individual and – although providing a foundation for performance consistency – is imperfectly translated into behavior or action. A third assumption is that MT should be viewed as a continuous concept whereby individuals may be more or less mentally tough, rather than mentally tough or not, or as comprising different profiles of MT (e.g., high in self-efficacy and optimism but low in perseverance); that is, individuals will typically score similarly on each of the key dimensions (cf. Gucciardi & Jones, 2012). This view of MT is also consistent with resource-based perspectives of human adaptation, coping and well-being (Hobfoll, 1989) in which it is said that having higher levels of one personal resource such as self-efficacy is typically associated with higher levels of other resources such as optimism, resilience, and hope (Stajkovic, 2006). Finally, aligned with leading
theories of stress (Hobfoll, 1989, 2002; Lazarus & Folkman, 1984), the key MT facets pertain
to everyday hassles (e.g., mundane pressures such as balancing work or study with other
aspects of one’s life, maintaining quality relationships with others) as well as major life
events that pose a significant threat to one’s normal level of performance or goal attainment
consistency (e.g., failed a course, major illness, death of a loved one). Therefore, we propose
that MT is fundamentally important for striving (e.g., working towards self-defined goals or
objectives), surviving (e.g., dealing with challenges, demands, or adversities), and thriving
(e.g., sustaining high levels of performance, experiencing vitality and learning or
performance gains).

**Dimensionality of Mental Toughness**

Whereas a definition of a concept ought to be described in relatively general terms,
the nature and number of dimensions should be encapsulated by a corresponding model
(Bacharach, 1989). We adopted two primary criteria to determine whether a variable was
included in a model of MT: (i) a personal resource or aspect of the self that is positively
linked with subjective or objective performance across at least two different stressful
achievement contexts (e.g., education, sport); and (ii) collectively, the personal resources
encompassed cognitions, emotions, and behaviors. Indeed, scholars (e.g., Covington &
Omelich, 1988; Raymond, 2009) have highlighted the importance of considering the dynamic
interaction among cognitive, emotional, and motivational domains when conceptualizing
processes related to achievement dynamics. We identified seven core personal resources from
an extensive review of the existing MT literature and our unpublished interviews with non-
sport performers that have the strongest empirical support in terms of their link with
performance and goal attainment or progress. The majority of these facets resemble existing
concepts (e.g., hope, emotion regulation, optimism, self-efficacy) that are generally supported
by well-established theories, frameworks or models (see Table 2).
Our research will also test a key feature of MT conceptualization that has yet to be directly examined; specifically, is MT best conceptualized as a unidimensional or multidimensional concept? When considered alongside previous work on MT (e.g., Clough et al., 2002; Jones et al., 2002), our guiding theoretical framework (Hobfoll, 1989, 2002), and related research (e.g., Luthans, Avolio, Avery, & Norman, 2007) and theory (Stajkovic, 2006), the various dimensions we have identified from our literature review and unpublished interviews with non-sport performers support a multidimensional perspective. However, it is important to recognize that much of the early work on MT adopted an exploratory, qualitative approach with sport performers and therefore limits the extent to which these findings generalize across samples and achievement contexts. Although exceptions do exist, equivocal findings have been revealed when a multidimensional conceptualization of MT has been psychometrically examined across diverse samples (cf. Gucciardi et al., 2012; Perry, Clough, Crust, Earle, & Nicholls, 2013). Perhaps most important for MT concept development, the notion of a multidimensional conceptualization appears to have been accepted by scholars with little research conducted to directly test this assumption against a unidimensional model.

Hypothesis 1: MT is best conceptualized as a multidimensional concept when compared with a unidimensional representation

Stress, Coping and Adversity

Resources including objects, conditions and personal characteristics play a central role in most models of the stress process. Within the context of a transactional perspective (Lazarus & Folkman, 1984), for example, stress is said to occur when individuals perceive events or situations in their environment to be taxing or exceeding their resources. Similarly, the interaction between demands and resources is a core assumption of the Jobs Demands-Resources model (Demerouti, Bakker, Nachreiner, & Schaufeli, 2001), such that stress or burnout develops when demands are high and resources are low. A central tenet of the
conservation of resources theory is the notion that people strive to obtain, preserve, and protect resources (Hobfoll, 1989, 2002). In clarifying its conceptualization, therefore, we propose that MT could be a useful organizing framework or “resource caravan” (Hobfoll, 2002) for those core personal resources central to the coping process in which individuals managed the internal and external stressors of the person-environment relationship (Lazarus & Folkman, 1984).

Resources are also an integral feature of resilience models (e.g., Masten, 2011; Windle, 2012) that deal with significant adversities including discrete experiences (e.g., parental psychopathology, community violence), cumulative indices (e.g., tallies of life adversities over time), and acute trauma and chronic life difficulties (e.g., sexual abuse, neighborhood disorganization) (Obradović, Schaffer, & Masten, 2012). Although there remains considerable debate regarding a formal definition, common themes among most contemporary conceptualizations reveal that resilience encapsulates a dynamic process whereby one regains or sustains relatively stable, healthy levels of psychological and physical functioning, or experiences positive adaptation following exposure to significant adversity (Masten, 2011; Windle, 2011). Individual (e.g., psychological resources, biological factors), community (e.g., social support) and societal (e.g., health and social services) protective factors are central to this process of recovering from or adjusting to adversity (Masten, 2011; Windle, 2012). Thus, conceptualized as a collection of core personal resources, MT can be conceived as an important individual protective factor within the resilience process but should not be conceptually equated with this phenomenon.

Also central to this process perspective of resilience is that it “is not necessarily about superior functioning or flourishing, it is about doing okay, or better than could be expected, given the individual circumstances” (Windle, 2012, p. 159). In other words, resilience is best conceptualized as the process by which one bounces back or recovers from such major
assaults, which is consistent with the original and basic meaning of the word\(^2\) (i.e., “spring back into shape” and “recover or adjust”). This “bounce back” conceptualization of resilience facilitates an important conceptual distinction between returning to normal levels of functioning (i.e., resilience) and moving towards a superior level of functioning or thriving following an adversity (Carver, 1998). Thriving is defined as “the psychological state in which individuals experience both a sense of vitality and a sense of learning” (Spreitzer, Sutcliffe, Dutton, Grant, & Sonenshein, 2005, p. 538). As an internal marker of individual growth and upward trajectory, thriving is considered state-like and malleable, and is therefore contingent upon the situation or task (Spreitzer & Porath, in press). With its focus on active, intentional engagement in the process of attaining and sustaining high levels of performance and/or goal attainment, MT should also play an important role in thriving.

Hypothesis 2: MT is inversely related with perceived stress, and positively associated with performance, goal attainment, and thriving

On the Traitness of Mental Toughness

An unresolved yet substantively important issue is whether MT is best conceptualized as a dispositional, trait-like or situational, state-like construct. It has been argued that “most psychological constructs vary along a continuum of stability or what [is referred to] as traitness” (Kenny & Zautra, 2001, p. 243). Some scholars (e.g., Clough et al., 2002; Hardy et al., in press) have proposed that MT represents a relatively stable dispositional trait, although some authors (e.g., Harmison, 2011) have argued for a state-like conceptualization. Despite these conflicting viewpoints, little empirical research has been conducted to directly examine this key feature of MT for concept development. Initial work in this area that gathered key stakeholders’ retrospective viewpoints provided indirect support for a state-like conceptualization; in other words, MT appears to be open to development and can both vary

\(^2\) We consulted both the Merriam-Webster and Oxford English online dictionaries.
and have enduring properties across situations or time (e.g., Coulter et al., 2010; Jones et al., 2002). Subsequent research supported these claims by showing that MT could be enhanced among a sample of adolescent footballers via a psychological skills training intervention (Gucciardi, Gordon, & Dimmock, 2009b). Contrary to these initial reports, however, researchers have revealed an almost perfect test-retest relationship of informant-rated mentally behavior over three weeks ($r = .96$) among a sample of 59 tertiary sport science students (Hardy et al., in press).

The equivocal nature of the available research means that uncertainty remains with regard to the traitness of MT. In particular, methodological features of this research limit our confidence in the findings. For example, the small sample sizes and focus on sport participants limited the extent to which these findings generalize to larger or more diverse populations. With regard to the temporality of these designs, retrospectively recalling mental processes may not be entirely accurate (Stone et al., 1998) and encompassing only two points across short time frames such as three weeks does little to capture the stability and dynamic nature of key concepts and their interrelationships (Chan, 1998). Also lacking in this previous research is guidance from established theory in terms of the rationale for expectations regarding the dispositional nature of psychological constructs. In this regard, Coulter, Mallett, Singer, and Gucciardi (2013) offered an alternative and broader view of mental toughness within the theoretical confines of contemporary personality psychology. They argued that mental toughness may be more appropriately understood across and within the multiple layers of personality as guided by an integrative personality framework (McAdams & Pals, 2006). McAdams and Pals (2006) conceptualized personality as encompassing key individual difference variables situated across diverse layers of understanding, including dispositional traits (broad dimensions that evidence consistency across situations and over time), characteristic adaptations (contextualized expressions of traits which are activated by
contextual or social factors), and life narratives (internalized and evolving psychosocial
construction of one’s identity). As we have conceptualized MT as a caravan of personal
resources pertinent to the coping process, we expected there to be some variation and stability
in this concept across situations and time as individuals assess whether or not the
characteristics of the event or situation exceeds their personal resources (Hobfoll, 1989, 2002;
see also, Harmison, 2011).

Hypothesis 3: The variance in MT will be accounted for by both within- (i.e., state-like) and
between-person (i.e., trait-like) differences

Overview of the Present Research

In light of the varied multidimensional conceptualizations and accompanying
fragmented literature base, together with the largely myopic research focus on sport, we
began our research by investigating the content universe (Study 1) and factor structure of MT
across three achievement contexts (Study 2). Our focus was on sport, educational, and
workplace contexts because they each cover a variety of different stressors and adversities
that individuals must successfully negotiate to perform well. For example, in educational
contexts, one’s performances only affect oneself, whereas in sport and business one’s
performances affect oneself and one’s team (e.g., teammates, peers, supervisors). We then
tested central features of our conceptualization of MT including its role in the stress-
performance relationship (Study 3); predictive validity in terms of goal progress, thriving and
psychological health over a 10-week period, as well as its traitness (Study 4); and its
predictive and incremental validity for selection testing in the military context (Study 5). We
obtained approval from an institutional human ethics committee for each of the studies prior
to data collection. In all studies, all participants were assured of confidentiality and
anonymity in responses, and informed of the voluntary nature of their participation and their
right to withdraw participation at any time before or after obtaining their consent.
Study 1 – Item Generation

We first created a pool of items that were designed to reflect the key dimensions of MT detailed in Table 2 but also hold meaning for our target populations (e.g., “I am determined to reach my full potential” and “I interpret adversity as a positive challenge”). Items were designed so that they could be adapted across achievement contexts such as sport, education, military and work. Before piloting a questionnaire with the target populations, it is important to obtain experts’ assessments of the conceptual model and item content to ensure that one is sampling all content that is potentially relevant to the construct of interest (DeVellis, 2003). It also important to capture the views of the “end users” by piloting new surveys with a representative sample of the target populations (Vogt, King, & King, 2004).

Thus, the aims of Study 1 were to (a) examine experts’ views and opinions on the quality of our conceptual model of MT, and (b) to create and provide evidence for the face and content validity of a pool of items designed to capture this conceptualization.

Method

Participants

A total sample of 30 academic experts and representatives of the target populations participated in this study. We purposefully sampled nine researchers (3 females) with expertise in the psychology of performance and scale development from higher education institutions on four continents (i.e., Australia, New Zealand, United Kingdom, USA) via email invitation. Representatives of the target population were conveniently sampled from an Australian institution and via the authors’ personal networks: nine tertiary students (3 females), six athletes (2 females), three coaches (all male), and four businesspeople (2 females). All academics, athletes, coaches, and businesspeople were experienced in their achievement context (i.e., experience >10 years); tertiary students were nearing the end of a
Masters Coursework degree in applied psychology \( (n = 5) \) or in their first year of an undergraduate course in psychology \( (n = 4) \).

**Procedures**

Focus groups were conducted with representatives of each achievement context (i.e., tertiary education, sport, business). Participants were provided with descriptions of the key dimensions and items designed to measure each characteristic and asked to identify whether the item corresponds to the dimension on which it is hypothesized to correspond to, as well as detail any issues with the item (e.g., length, readability). Subsequently, academic experts were invited by email to provide feedback using an online questionnaire developed for this study. Experts were asked to rate (i.e., 5 point Likert scale: 1 = *poor*, 3 = *good*, 5 = *excellent*) and comment (i.e., open-ended responses) on the breadth of the key dimensions and the definitions for each of the subscales, as well as rate the adequacy of each of the items for capturing its hypothesized component of MT (i.e., 5 point Likert scale: 1 = *poor match*, 3 = *good match*, 5 = *excellent match*).

**Results and Discussion**

We developed 70 items designed to assess the key dimensions of our conceptualization of MT (see Table 2) and obtained stakeholder’s perspectives on the adequacy of these items and our conceptual model. Representatives of the target populations deemed all but seven items as corresponding to the MT facet it was intended to assess and were therefore eliminated. Of the remaining 63 items, nine were reworded to improve their clarity and applicability across different achievement contexts. The academic experts believed that our conceptual model was very good both in terms of breadth of the key dimensions \( (M = 4.33) \) and quality of characteristic descriptions \( (M = 4.22) \). The item pool \( (n = 63) \) was also rated by our academic experts and subsequently used to calculate a Content Validity Index (CVI; Lynn, 1986) for each item; that is, the number of experts who gave a
rating of 3 or above divided by 9 (i.e., the total number of experts). We used the CVI (i.e., > .80 deemed acceptable; Lynn, 1986) and open-ended responses to inform final decisions about whether to retain, eliminate, or revise an item. Two items with a CVI of .78 (7/9) were deemed unacceptable and therefore removed from the item pool. Although the remainder of the item pool evidenced adequate CVIs, the wordings of 14 items were modified slightly to improve clarity and reduce overlap with other key dimensions. This process led to 61 items being retained for inclusion in the initial survey.

Study 2 – Dimensionality of Mental Toughness

We administered the preliminary pool of items to participants from our target populations, namely students, athletes, and employees. The initial item pool developed in Study 1 was over-inclusive so we focused on refining these items down to a final, shorter version by selecting the best items according to the obtained statistical or psychometric properties of the items (Marsh, Hau, Balla, & Grayson, 1998) alongside conceptual criteria (i.e., content validity). Within-network evidence (e.g., factorial validity, reliability) represents “a logical prerequisite” (Marsh, 1997, p. 28) for instruments prior to examining more substantive issues such as locating the construct in a broader conceptual space. Consistent with our first hypothesis, we expected a multidimensional model encompassing seven first-order factors (see Table 2) and one second-order factor of MT to gain support.

Method

Participants

Convenience samples from education, sport, and workplace contexts in Australia were recruited to participate.

Calibration sample. A total of 418 performers participated: 136 athletes (57 males, 76 females; $M_{age} = 27.21$, $SD_{age} = 10.36$), 137 students (58 males, 78 females; $M_{age} = 19.99$, $SD_{age} = 10.98$) for the TMT and 145 employees (76 males, 75 females; $M_{age} = 30.75$, $SD_{age} = 10.08$) for the SCT.
Cross-validation samples. Three independent samples were recruited: 500 students (176 males, 323 females; $M_{age} = 20.43, SD_{age} = 4.70$), 427 athletes (167 males, 269 females; $M_{age} = 27.21, SD_{age} = 10.34$), and 550 “white collar” workers (304 males, 243 females; $M_{age} = 48.14, SD_{age} = 10.27$). Age and gender were not voluntarily reported by some participants.

Measure and Procedures

Tertiary students were recruited from a large Australian university and included individuals enrolled in undergraduate courses in psychology, sport science, and commerce; psychology and sport science students received course credit as part of an established research participation scheme, whereas commerce students received a lecture on the topic including an overview of their findings. Athletes were recruited via national sporting organizations, which included a variety of team (e.g., netball, waterpolo, Australian football) and individual sports (e.g., triathlon, swimming, golf) across all levels of competition, including amateur, semi-professional and professional competitions. The lead author obtained gatekeeper approval before invitations were distributed by key personnel (e.g., Manager of Science and Medicine, CEO) to athletes. Employees in full-time employment were recruited via the Australian Institute of Management, a not for profit entity designed to promote the advancement of education and learning in the fields of management and leadership for commerce, industry and government. An email invitation was sent to members by the CEO on behalf of the research.

Informed consent and survey questions were completed online using a secure data collection site. The survey contained the 61 items developed in Study 1 to capture the hypothesized components of MT. Participants were instructed to indicate how true each statement was an indication of how they typically think, feel, and behave as a performer in
their context (e.g., as an athlete). We adopted a 7-point Likert scale (1 = false, 100% of the time to 7 = true, 100% of the time) for the MT items, as this response format has been shown to optimize reliability, validity, discriminating power, and respondent preferences (Preston & Colman, 2000). Participants were recruited from a large tertiary institution, via national sporting organizations, and organizations in the services sector (e.g., education, health care).

Data Analysis

Preliminary analyses. The data were first examined for missing values, violations against assumptions of univariate normality, and both univariate (i.e., z score > ± 3.29) and multivariate outliers (i.e., using a p < .001 criterion for Mahalanobis D²) in IBM SPSS 20. With regard to normality, skewness and kurtosis values that exceed 2 and 7, respectively, indicate a non-normal distribution at the univariate level which can affect goodness of fit statistics (Tabachnick & Fidell, 2007).

Primary analyses. The factorial validity of the hypothesized MT model was tested with confirmatory factor analysis (CFA) in Mplus 7 (Muthén & Muthén, 2012) using a robust maximum likelihood estimator (MLR). The χ² goodness-of-fit statistic, comparative fit index (CFI), Tucker-Lewis index (TLI), standardized root mean square residual (SRMR), and root mean square error of approximation (RMSEA) were used to evaluated the fit of the models. Browne and Cudeck’s (1992) guidelines were adopted as indicators of good fit (CFI, and TLI ≥ .90, and SRMR and RMSEA scores ≤ .08).

In keeping with the three-stage strategy advocated by Jöreskog (1993) and employed by others (e.g., Lonsdale, Hodge, & Rose, 2008), we initially examined one-factor congeneric models whereby CFA is performed on individual subscales using the calibration sample. We next performed a series of two-factor CFAs in which each dimension was paired with each of the other subscales. In the third and final stage, we tested the model in its entirety. Items were considered for deletion if modification indices suggested that the error term of an item
correlated with that of another item, an item had a low factor loading (<.50), or modification indices suggested that an item cross-loaded on an unintended latent variable (Ford, MacCallum, & Tait, 1986). Conceptual issues (e.g., construct breadth) were also considered alongside these statistical criteria. The resulting model was subsequently tested on the cross-validation samples.

In all analyses, the discriminant validity of the latent factors was assessed by examining the 95% confidence intervals of the interfactor correlations. Discriminant validity is supported when the 95% confidence interval of the interfactor correlations does not encompass ± 1.0 (Anderson & Gerbing, 1988). With regard to convergent validity, we examined standardized solutions to evaluate the significance and strength of parameter estimates. Standardized factor loadings are interpreted using Comrey and Lee’s (1992) recommendations (i.e., > .71 = excellent; > .63 = very good; > .55 = good; > .45 = fair; < .32 = poor). Finally, a composite reliability coefficient (ρ; Raykov, 1997) was calculated within a structural equation modeling framework to estimate the level of internal reliability for each factor (i.e., acceptable levels of reliability > .70; Nunnally & Bernstein, 1994).

Results and Discussion

Preliminary Analyses

The data were examined for missing values prior to the main analyses. When less than 5% of data points are randomly missing, almost any procedure for handling missing values yields similar results (Tabachnick & Fidell, 2007). Missing data (1.2% on all data points) did not relate to any of the demographic variables and were handled in IBM SPSS 20 using the expectation-maximization method. Data screening procedures did not reveal any univariate or multivariate outliers. However, item-level analyses revealed that some items were negatively skewed. Thus, we employed a robust maximum likelihood estimator (MLR) in all subsequent modeling, which produces standard errors and tests of fit that are robust in relation to non-
normality of observations and the use of categorical variables when there are at least four or more response categories (Beauducel & Herzberg, 2006).

Primary Analyses

In stage one, the series of single subscale CFA resulted in the deletion of 19 items according to both empirical (e.g., low factor loadings, high cross-loadings) and conceptual (e.g., items with similar meanings) considerations. Each of the one-factor congeneric models displayed excellent fit with the data according to our multiple criteria. In stage two, an additional 21 items were deleted based on the results of a series of CFAs involving two latent factors. Finally, as detailed in Table 3, the resultant 21-item, seven-factor model evidenced good fit with the data in all samples as both a lower-order (i.e., correlated latent factors) and higher-order model. Composite reliabilities ($\rho$) of the seven latent factors were adequate in the calibration ($\rho = .74$ to $.89$), athlete ($\rho = .75$ to $.90$), tertiary student ($\rho = .71$ to $.84$), and workplace samples ($\rho = .70$ to $.86$). An inspection of Table 4 reveals that we achieved good (i.e., $>.55$) factor loadings for each item, with the average loadings of all 21 items considered excellent (i.e., $>.71$) across all four samples. Despite the encouraging findings of the multidimensional models regarding fit indices, internal reliability, and factor loadings, there were several instances of inadmissible solutions resulting from a “not positive definite” covariance matrix. A non-positive definite matrix may be due to small sample size (e.g., < 300), model over fitting, empirical under-identification, or model misidentification (Wothke, 1993). Model misidentification appeared to be the most plausible explanation for the nonpositive definite matrix in our case$^3$. Additionally, an inspection of the 95% confidence intervals of the latent factor correlations and higher-order coefficients revealed that several relationships encompassed 1 thereby suggesting a lack of discriminant validity.

$^3$ We subjected both the lower-order and higher-order seven factor models to a CFA with a combined sample involving the athletes, students, and employees (N=1492). In both instances, the latent covariance matrix was not positive definite thereby highlighting the inadequacy of these models.

This article is protected by copyright. All rights reserved.
Our approach to establishing initial reliability and validity evidence for our hypothesized higher-order, multidimensional model (i.e., seven lower-order factors) of MT was unsuccessful. Although each of the lower-level facets are backed by an extensive history of research and theory (see Table 2), and were supported in our three samples using one-factor congeneric models, our findings suggested that they do not necessarily sum to a coherent whole. The large correlations observed across three independent samples suggested considerable empirical redundancy between the key dimensions of MT, despite the conceptual distinctions proposed by scholars (e.g., Luthans, Avolio et al., 2007; Stajkovic, 2006). High correlations have also been observed elsewhere when self-efficacy, optimism, hope, and resilience have been examined as a resource caravan (e.g., Caza, Bagozzi, Woolley, Levy, & Caza, 2010; Luthans, Norman, Avolio, & Avey, 2008).

The majority of previous research which has sought to conceptualize MT has adopted a qualitative approach in which a small number of key stakeholders’ (e.g., athletes, coaches) have retrospectively recalled their experiences to generate an understanding of this concept (for reviews, see Gucciardi & Gordon, 2011). Thus, it was considered important that these initial findings were tested on large samples across multiple achievement contexts. Contrary to expectations regarding a multidimensional conceptualization of MT (cf. Clough et al., 2002; Coulter et al., 2010; Jones et al., 2002), the findings of this study indicated that the (often subtle) conceptual distinctions among these key personal resources made by scholars are not readily made by individuals from diverse achievement contexts thereby suggesting that a unidimensional model may be more appropriate than a multidimensional representation.

Substantively, when conceptualized as a higher-order concept, MT must be inferred on the basis of factor analyses of scales that capture the key dimensions. This indirect approach to conceptualizing higher-order concepts can lead to confusion over whether the
multifaceted concept is a latent or aggregate construct (Carver, 1989). Thus, we subsequently examined the utility of a direct approach for measuring the underlying concept itself rather than the key dimensions of MT. A direct approach is also consistent with the emerging trend toward short measures of psychological concepts (e.g., Schulenberg, Schnetzer, & Buchanan, 2011; West, Dyrbye, Satele, Sloan, & Shanafelt, 2012). Items were selected from the 21 items displayed in Table 4 according to a combination of empirical (e.g., size of standardized loading, minimal cross-loadings) and conceptual (e.g., ratings from expert judges in Study 1, adequately captured the content domain of each facet) criteria. One item for each facet was retained, as well as an item to measure one’s capacity to deal with adversity so as to capture both everyday challenges and major distresses. The 8-item undimensional model evidenced excellent fit (see Table 3) and good-to-excellent factor loadings (see Table 5) across all four samples. Composite reliabilities for the unidimensional measure of MT were excellent ($\rho = .86$ to .89).

**Conclusion.** Aligned with our first research hypothesis, our failure to support the discriminant validity of the hypothesized key dimensions of MT in the multidimensional model is at odds with our expectation and raised the question as to whether this concept is undimensional rather than multidimensional as previously espoused by several scholars (e.g., Clough et al., 2002; Coulter et al., 2010; Jones et al., 2002). Subsequent analyses indicated that the 8-item, direct assessment of unidimensional MT fit the data very well, displayed strong factor loadings, and produced an internally reliable score across three independent samples of performers. The brevity of the direct approach may offer practical appeal not only in academic but also field settings. Despite these encouraging findings, this study was limited to an examination of the within-network properties of mental toughness. Thus, it is important to ascertain whether MT functions in a theoretically expected way with regard to key features.
of our working definition, namely perceived stress, performance, goal attainment, and
thriving.

Study 3 – Mental Toughness, Stress and Performance in the Workplace

An inverse relationship between stress and performance has been evidenced across
variety of achievement contexts, including the workplace (Gilboa, Shirom, Fried, & Cooper,
2008) and educational settings (Richardson, Abraham, & Bond, 2012). Drawing from the
transactional theory of stress (Lazarus & Folkman, 1984), a situation is appraised as stressful
when an individual concludes that demands imposed by or perceived within a given context
exceeds his or her available coping resources. Conceptualized as a personal capacity or
resource, MT should play a role in determining the extent to which individuals perceive
transactional experiences and their lives in general to be “unpredictable, uncontrollable, and
overloaded” (Cohen, Kamarck, & Mermelstein, 1983, p. 385). Substantively, therefore, we
proposed that the perceived stress concept should help explain the relationship between MT
and performance. In other words, MT is postulated to exert some of its influence on
performance through the distress and coping dimensions (i.e., mediation or indirect effect).

Method

Participants and Procedure

An email invitation containing the study information was distributed to personal
contacts of the research team, who subsequently disseminated the study information to their
colleagues; this snowball approach resulted in a convenience sample of 497 employees (275
male and 219 female, 3 did not report gender; $M_{age} = 47.98$, $SD = 9.95$). Employees were
“white collar” workers in full-time employment drawn from different organizations and
diverse ranks of the Australian services sector such as education, health care and finance.
Informed consent and survey questions were completed online using a secure data collection
site.
Measures

MT. The 8-item MTI developed in Study 2 was employed to measure MT, and was found to be internally reliable in the current sample ($\rho = .86$).

Perceived stress. The 10-item Perceived Stress Scale (Cohen et al., 1983) was used to assess an individual’s subjective appraisal of global stress during the last month (e.g., “felt nervous and stressed” and “found that you could not cope with all the things that you had to do”). Items were rated on a scale from 0 (never) to 4 (very often). A correlated two-factor model of perceived stress involving distress (6 items) and coping (3 items) dimensions was employed because this factor structure has received support across diverse samples (e.g., Lavoie & Douglas, 2012). The distress ($\rho = .87$) was found to be internally reliable in the current sample; however, coping ($\rho = .65$) was just below the minimum recommend level of .70 (Nunnally & Bernstein, 1994).

Performance. Each participant’s workplace performance over the last month was assessed by his or her supervisor using a 7-item measure (e.g., “Adequately completes assigned duties” and “Fulfills responsibilities specific in job description”) of in-role behaviors (Williams & Anderson, 1991). An informant-rated approach is consistent with recent research on mentally tough behaviors (Hardy et al., in press). Items were rated on a scale from 1 (strongly disagree) to 7 (strongly agree). Performance was found to be internally reliable in the current sample ($\rho = .86$).

Results and Discussion

We estimated and tested the significance of the indirect effect of MT on informant-rated performance via perceived stress using the bootstrapping technique and 5000 resamples (Preacher & Hayes, 2008). This method does not assume normality of the distribution of indirect effects and uses a resampling procedure to create confidence intervals for the indirect effect (Hayes, 2009). Bootstrapping allows for the examination of the total, direct and
indirect effects of the dependent variable on the independent variable. A significant mediation or indirect effect is observed when the 95% confidence interval (CI) that is obtained from the analysis does not include zero. These analyses were conducted within a structural equation modeling framework using Mplus 7 (Muthén & Muthén, 2012).

The fit statistics for the meditational model indicated acceptable fit with the data, \( \chi^2 (246) = 618.06, p < .001, \) CFI = .923, TLI = .914, SRMR = .045, RMSEA = .055 (90% CI = .050 to .061). Direct path coefficients from MT to distress (\( \beta = -.52, B = -.63, p < .001 \)), coping (\( \beta = .51, B = .44, p < .001 \)), and performance (\( \beta = .34, B = .41, p < .001 \)) were significant. Performance evidenced direct relationships with coping (\( \beta = .31, B = .44, p < .001 \)) and distress (\( \beta = -.15, B = -.15, p < .05 \)). The total indirect effect from MT to performance via distress and coping was significant (standardized indirect effect estimate = .24, \( p < .001, 95\% CI [.17, .31] \)). The specific indirect effects of MT to performance via distress (standardized indirect effect estimate = .08, \( p < .05, 95\% CI [.01, .15] \)) and coping (standardized indirect effect estimate = .16, \( p < .001, 95\% CI [.08, .24] \)) were significant, thereby supporting our expectation that MT would exert some of its influence on performance through the distress and coping dimensions. MT, distress, and coping explained 45% of the variance in performance, whereas MT explained 27% and 26% of the variance in distress and coping, respectively.

Consistent with our second research hypothesis, three key features of our conceptualization of MT received support in this study. First, MT has been commonly associated with high performance yet little evidence was offered to support this conclusion. As self-reported MT was directly associated with higher levels of supervisor-rated performance, the results of this study are among the first to provide some support for this theoretical proposition. Second, the findings provided initial validity evidence to support an inverse relationship between distress, and a positive association with coping dimensions of...
perceived stress. Third, the findings offer preliminary insight into the mechanisms by which MT has an effect on performance in the workplace. Employees with higher levels of MT are less distressed and better able to cope with the demands or challenges in their life, which in turn are associated with higher levels of performance. Consistent with our conceptualization, MT appears to play a role in influencing one’s perception of whether his or her personal resources are sufficient to cope with the person-environment interactions and therefore the extent to which an individual appraises challenges or demands as stressful (cf. Lazarus & Folkman, 1984).

**Study 4 – Mental Toughness and Key Outcomes in Education**

In this study, we sought to replicate and extend upon the results of Study 3 in several ways. First, we extended our focus to an educational context in which achievement is also a central feature. Second, consistent with our guiding definition and conceptualization, perceptions of goal progress and thriving (Spreitzer et al., 2005) were examined as hypothesized outcomes of MT. Third, owing to the prevalence of mental health issues among young people aged 16 to 34 years of age (e.g., approximately 25% in Australia; ABS, 2008), we also explored psychological health as a potential outcome of MT. Psychological health was conceptualized as consisting of the presence of positive symptoms (i.e., positive emotions) and the absence of negative symptoms (i.e., negative emotional states including depression, anxiety) (Keyes, 2003). Finally, we conducted a repeated measures design to explore prospective relationships among the study variables over the course of a university semester. Aligned with our third research hypothesis, obtaining repeated measurements of the same individual over time enables one to simultaneously model within-person (i.e., state) and between-person (i.e., trait) variability in study variables (Roesch et al., 2010) thereby permitting an examination of the traitness of MT. Such an approach enables an exploration of
an individual’s current MT relative to their usual level (i.e., within-person variability), as well as relative to other people’s MT (i.e., between-person variability).

Two research hypotheses underpinned this study. First, guided by our concept definition and theoretical model, we expected MT to be positively associated with goal progress, thriving, and our positive indicator (i.e., positive emotions) of psychological health, as well as inversely related with our negative indicator of psychological health (i.e., negative emotional states). Second, guided by both theory (Hobfoll, 1989, 2002) and research (e.g., Harmison, 2001; Kenny & Zautra, 2001), we hypothesized that the variance in MT would be accounted for by both within- (i.e., state-like) and between-person (i.e., trait-like) differences in MT.

Method

Participants and Procedure

A total of 203 undergraduate sport science students from an Australian university (92 male and 105 female; 6 participants did not specify gender; \(M_{\text{age}} = 20.69, SD = 2.62\)) participated in return for course credit. The lead author provided an overview of the study to all students in the first lecture of the semester. Students completed an online survey containing all study variables listed below every week for ten weeks over the course of a university semester. A reminder email containing the web link was sent every Friday afternoon to enrolled students who agreed to participate in the study. Participants completed the survey within the next 24 hours. Of the 203 participants registered for the study, three participants completed four surveys; four participants completed five surveys; 12 participants completed six surveys; 16 participants completed seven surveys; 29 participants completed eight surveys; 53 participants completed nine surveys; and 86 participants completed 10 surveys (\(M_{\text{cluster size}} = 8.79\)).

Measures
The weekly diary survey package contained measures of MT, goal progress, thriving, and psychological health. Owing to the repeated administration of the survey package, all scales with the exception of MT were reduced to achieve an approximate completion time of 5 minutes and therefore minimize study fatigue. Pilot data not reported here were collected from a larger sample of undergraduate students (N = 533) from a variety of courses (e.g., psychology, sport science, business) to guide empirical decisions for item selection (i.e., highest standardized factor loading) alongside theoretical criteria (i.e., adequate representation of the intended construct). For each measure, participants were instructed to indicate how much the statements applied to them as a university student over the past week (including dates). The ordering of the scales within the survey package was altered each week to minimize order effects.

**MT.** We employed the 8-item MTI developed in Study 2.

**Thriving.** Six items from the Thriving Scale (Porath, Spreitzer, Gibson, & Garnett, 2012) were employed to assess the cognitive (learning) and affective (vitality) components of thriving in the educational context. Sample items include “I continued to learn more and more as time went by” and “I had energy and spirit”. Items were rated on a scale from 1 (strongly disagree) to 5 (strongly agree).

**Goal progress.** Participants listed an academic and social goal which they intended to pursue over the course of the University semester at the first data collection point. The instructions for nominating personal goals were taken directly from previous research (Koestner, Otis, Powers, Pelletier, & Gagnon, 2008) and read as follows: “Personal goals are projects and concerns that people think about, plan for, carry out, and sometimes (though not always) complete or succeed at. They may be more or less difficult to implement; require only a few or a complex sequence of steps; represent different areas of a person’s life; and be more or less time consuming, attractive, or urgent. Please list the most important University
or study-related goal that you have for this coming semester. Now do the same for your most important social goal” (p. 235). At each follow-up survey, participants rated how much progress they had made toward each goal using a scale from 1 (none) to 9 (a great deal).

**Psychological health.** As health and illness are not considered the exact opposites of a single continuum (Keyes, 2003), it was important to assess psychological health both in terms of the presence of positive symptoms (i.e., positive emotions) and the absence of negative symptoms (i.e., negative emotional states). Positive emotions were measured with four items used in previous research (Tamir, John, Srivastava, & Gross, 2007). Items were rated on a scale from 1 (not at all) to 5 (all of the time). Sample items include “Happy / Pleased / Contented” and “Proud / Accomplished / Successful”. Negative emotional states were measured with six items from the Depression, Anxiety and Stress Scales (Lovibond & Lovibond, 1995). Sample items include “I found myself getting agitated” and “I felt I had nothing to look forward to”. Items were rated on a scale from 0 (did not apply to me at all) to 3 (applied to me very much/most of the time).

**Results and Discussion**

As the data represented a 2-level structure – observations (Level 1) nested within participants (Level 2) – multilevel structural equation modeling (MLSEM; Preacher, Zyphur, & Zhang, 2010) was employed to examine the study hypotheses. Multilevel models accommodate the hierarchical nature of data characterized by non-independence (i.e., two observations from the same person are likely to be more similar than two observations from different persons) by simultaneously estimating the variability at each level of the data hierarchy. With the exception of goal progress, all constructs were modeled as latent variables with items representing observed indicators. The intercepts or means of Level 1 variables were allowed to vary across Level 2 units (i.e., people). All analyses were conducted within Mplus 7 (Muthén & Muthén, 2012) using the MLR estimator to
accommodate for missing data and unbalanced cluster sizes (i.e., number of observations for each person). Composite reliability estimates ($\omega$) of the study variables were computed within a multilevel CFA framework (Geldhof, Preacher, & Zyphur, in press).

**Null model.** We first calculated the intraclass correlation (ICC) for each of the observed variables (i.e., item indicators for latent constructs, goal progress score) to examine the between-person variability in the study variables. The ICC can range from 0 to 1, with values close to zero (e.g., .05) suggesting that multilevel modeling may not be required (Dyer, Hanges, & Hall, 2005). At least 38% of the variance in the study variables was associated with between-person differences. With regard to MT, on average, 44% of the total variance is due to between-person differences (ICCs ranged from .38 to .50). The remaining 56% of the variance is due to the within-person variability across the 10 weeks of the study period. These findings supported our expectation that a comparable amount of variance would be accounted for by both within- (i.e., state-like) and between-person (i.e., trait-like) differences in MT. Consistent with a person-situation interaction perspective of MT (e.g., Harmison, 2011), these findings provide preliminary support for the idea that some individuals bring a dispositional aspect of MT to their interactions with the environment.

**Multilevel scale reliability.** Cronbach’s alpha, which assumes a single-level structure in one’s data, is typically reported to support scale reliability in research that encompasses multilevel data (Geldhof et al, in press). Drawing from recent guidelines (Geldhof et al., in press), we conducted separate multilevel CFAs to account for the clustered nature of our data and simultaneously assess scale reliability at both the within and between levels of analysis:

MT ($\omega_{\text{within}} = .80; \omega_{\text{between}} = .98$), thriving ($\omega_{\text{within}} = .78; \omega_{\text{between}} = .96$), negative emotional states ($\omega_{\text{within}} = .69; \omega_{\text{between}} = .91$) and positive emotions ($\omega_{\text{within}} = .80; \omega_{\text{between}} = .94$).

**Multilevel structural model.** We analyzed a full structural model in which MT served as a predictor of psychological health, thriving, and goal progress. The fit statistics
indicated acceptable fit with the data, $\chi^2(572) = 1772.37$, $p < .001$, CFI = .925, TLI = .915, SRMR$_{\text{within}} = .035$, SRMR$_{\text{between}} = .089$, RMSEA = .034. An overview of the parameter estimates for the multilevel structural model is detailed in Table 6. All of the direct relationships between MT and the outcome variables were statistically significant and consistent with our expectations. MT emerged as a statistically significant predictor of negative emotional states ($R^2_{\text{within}} = 10\%$; $R^2_{\text{between}} = 21\%$), positive emotions ($R^2_{\text{within}} = 33\%$; $R^2_{\text{between}} = 75\%$), thriving ($R^2_{\text{within}} = 41\%$; $R^2_{\text{between}} = 87\%$), and academic ($R^2_{\text{within}} = 15\%$; $R^2_{\text{between}} = 53\%$) and social goal progress ($R^2_{\text{within}} = 3\%$; $R^2_{\text{between}} = 45\%$) at both levels of analysis. The strength of the relationship between MT and the outcome variables were stronger at the between-person level when compared with the within-person level. These data also provided further evidence for our third hypothesis because the direct relationships between MT and all five outcome variables were significant at both levels of analysis. Specifically, MT appears to play an important role in understanding goal progress, thriving, and psychological health both in terms of variation within a person (e.g., over time) and between people.

Study 5 – Mental Toughness and Selection Testing in the Military

In this study, we tested the predictive and incremental validity of MT for selection testing in a military context using a naturalistic design. Aligned with our guiding conceptualization of MT, selection testing of already enlisted military personnel was chosen as an appropriate setting in which to examine these aspects of validity because it is a highly stressful course that requires candidates to sustain high levels of performance over an extended period of time (i.e., 6 weeks) across multiple and varied assessment tasks. Thus, this study was designed to provide an insight into whether or not MT was important for sustaining high levels of performance despite stress or adversity. We considered it important to test our new unidimensional conceptualization of MT against the popular 4Cs.

This article is protected by copyright. All rights reserved.
multidimensional model (i.e., control, commitment, challenge, confidence; Clough et al., 2002). In other words, what is the value of this new unidimensional conceptualization of MT above and beyond the existing multidimensional perspective of MT? Because the psychometric properties of the MTQ48, which is designed to capture the 4Cs model of MT, have been shown to be potentially problematic (Gucciardi et al., 2012), we employed established measures of the four key components of this hypothesized conceptualization of MT. We expected MT to predict success in the selection test, that is, contribute to whether one failed or passed the test, as well as retain its significance when control, commitment, challenge, efficacy were included as predictors of performance.

Participants and Procedure

A total of 115 male candidates aged 20 to 41 years ($M_{\text{age}} = 27.16$, $SD = 4.09$) with between one and 17 years of service ($M_{\text{years}} = 6.50$, $SD = 3.38$) in the Australian Defence Force participated in this study. These individuals voluntarily signed up to complete a six-week selection test for entry into the Special Forces unit. Participants completed a multisection survey containing all study variables listed below in the measures section prior to completing the selection test. Candidates completed the survey package on the first night of the course immediately following a briefing about the study by a research team member. To encourage honest responding, candidates were informed that the survey was for research purposes only and that their responses would be kept confidential and not influence their course outcome. No time limit was imposed on candidates for filling-in the survey, and return of completed surveys was taken as informed consent.

The selection test consisted of a six-week selection course specifically designed to assess each candidate’s suitability for entry into elite military training (i.e., Special Forces). The selection course comprised a range of individual and team activities that were both physically and mentally demanding, while being relevant to the military context (e.g.,
navigating between waypoints, planning a mission, firing weapons, carrying heavy loads for extended periods after minimal sleep). Candidate performance was continuously monitored and assessed throughout the course by instructional staff. Candidates could be removed from the course at any stage for not meeting the required standards or due to medical reasons (e.g., physical injury). Alternatively, candidates could elect to voluntarily withdraw from the course at any stage. Aligned with the highly stressful and challenging nature of the selection test, no feedback was given to candidates about their performance by instructional staff during the course. Each candidate’s final outcome on the selection test (i.e., pass / fail) was obtained by the fifth author with permission from a staff member involved in running the selection course; in total, 50 out of 115 candidates (43%) passed the course.

Measures

MT. The 8-item MTI developed in Study 2 was employed to measure MT, and was found to be internally reliable in the current sample (α = .84).

Hardiness. The 15-item Norwegian Dispositional Resilience Scale (Hystad, Eid, Johnsen, Laberg, & Bartone, 2010) to measure hardiness. This scale has three 5-item subscales designed to assess control (e.g., “How things go in my life depends on my own actions”), commitment (e.g., “I really look forward to my work activities”) and challenge (e.g., “I enjoy the challenge when I have to do more than one thing at a time”). Items were rated on a scale from 0 (not at all true) to 3 (completely true). Scale scores were computed by reverse scoring negatively keyed items and averaging across items, with higher scores indicating higher levels of each facet of hardiness. In this sample, Cronbach’s alpha for total hardiness (α = .73) and the commitment factor (α = .71) were good, whereas the control (α = .62) and challenge (α = .64) facets were below the minimum recommend level of .70 (Nunnally & Bernstein, 1994).
**Self-efficacy.** The 8-item New General Self-Efficacy Scale (Chen et al., 2001) was employed to assess an individual’s overall perceived efficacy or ability across different tasks and situations (e.g., “Compared to other people, I can do most tasks very well” and “In general, I think I can obtain outcomes that are important to me”). Items were rated on a scale from 1 (strongly disagree) to 5 (strongly agree). A total self-efficacy score was created by averaging the 8 items, with higher scores indicating higher levels of perceived efficacy. Self-efficacy was found to be internally reliable in the current sample ($\alpha = .88$).

**Selection test outcome.** Performance was coded as a dichotomous variable, where failure = 0 and pass = 1.

**Results and Discussion**

Missing data represented less than 0.5% of all data points and was therefore imputed using expectation-maximization method in IBM SPSS 20. Subscale level skewness (-.68 to 0) and kurtosis (-.89 to 1.15) estimates were acceptable, and none of the data violated assumptions of multivariate normality (Mahalanobis distance at $p < .001$) (Tabachnick & Fidell, 2007). However, three participants violated assumptions of univariate normality for the commitment ($n = 2$) and control subscales ($n = 2$); the removal of these participants did not alter the outcomes of the main analyses. MT was positively associated with the three hardiness facets ($r_{\text{commitment}} = .42; r_{\text{control}} = .44; r_{\text{challenge}} = .34$) and self-efficacy ($r = .70$).

Similar relationships were observed between self-efficacy and the hardiness components ($r_{\text{commitment}} = .37; r_{\text{control}} = .39; r_{\text{challenge}} = .34$). Hardiness subscales were positively associated with each other ($r = .20$ to .30). Nevertheless, the correlations need to be considered with caution given the low reliability estimates for the control and challenge subscales.

To examine the primary research question, we employed logistic regression in Mplus 7 (Muthén & Muthén, 2012) using a robust maximum likelihood estimator (MLR). The three hardiness facets of commitment ($B = .00, SE = .55, p = .99, \text{Odds Ratio } [OR] = 1.00$), control...
and self-efficacy (B = -.35, SE = .69, p = .62, OR = .72) did not emerge as a significant contributor to the prediction of selection test outcome; however, MT significantly predicted the performance outcome (B = 1.25, SE = .63, p < .05, OR = 3.48). Therefore, consistent with expectations, we found that MT was significantly associated with the successful completion of a rigorous military selection test, even when hardiness and self-efficacy are considered. These findings provide further evidence for the importance of MT for high performance and extend the data obtained with employees in Study 3. Specifically, as the selection test is conducted over six weeks and requires candidates to successfully complete multiple and varied assessment tasks, the findings of this naturalistic study provide initial evidence to support the importance of MT for sustaining performance over an extended period of time.

General Discussion

Despite its pervasiveness among both popular press and scholarly literature over the past decade, MT has suffered from both conceptual and methodological concerns thereby limiting its usefulness as a psychological concept. In an attempt to progress this area of research, we offered an alternative conceptualization of MT that drew from existing research (for reviews, see Gucciardi & Gordon, 2011) and related theory (Hobfoll, 2002; Lazarus & Folkman, 1984), and conducted a series of studies aimed at validating the concept. Of particular interest were fundamental issues pertaining to the dimensionality, nomological network, and “traitness” of MT. Collectively, this series of studies has offered three key substantive and methodological contributions to clarifying the scientific understanding of MT.

The first key contribution of this research pertains to the dimensionality of MT. Previous research (e.g., Clough et al., 2002; Coulter et al., 2010; Jones et al., 2002) and guiding theory (Hobfoll, 2002) led us to expect that MT would be best conceptualized as a
multidimensional concept when compared with a unidimensional structure. In testing these
two competing conceptual representations in Study 2, our expectation that MT represented a
higher-order concept that accounts for the shared variance or commonality among several
important personal resources was not supported. Instead, we identified a considerable degree
of empirical overlap or lack of discriminant validity between the key personal resources, with
a unidimensional model identified as an adequate fit with the data. The lack of discriminant
validity among the key dimensions of MT was also evident in the lower-order model in
which the attributes were allowed to freely correlate with each other. These findings are in
direct contrast to the commonly accepted multidimensional view of MT. Recent research
(Gucciardi & Jones 2012) and guiding theory (Hobfoll, 1989) offers a platform upon which
to understand the considerable empirical overlap between the MT facets; specifically, it is
said that having higher levels of one personal resource is associated with higher levels of
other resources (Stajkovic, 2006). Collectively, our findings suggested that the interpretations
regarding the dimensionality of MT scholars have made of performers’ perceptions of this
construct from qualitative research may not be entirely accurate, including our own early
work, and therefore require reconsideration and examination in future research.

The second key contribution of this research relates to the nomological network of
MT. Our conceptualization of MT as a personal capacity to produce consistently high levels
of subjective (e.g., goal progress) or objective performance (e.g., sales, race time, GPA)
despite everyday challenges and stressors as well as significant adversities received support
across several studies. First, in Study 3, we confirmed our substantive expectation of an
association with perceived stress in that MT was inversely related with the distress
dimension, and positively associated with the coping dimension. These findings align with
our expectation that MT plays an important role in determining how individuals perceive the
demands or challenges they encounter during their transactions with the environment
Specifically, individuals with higher levels of MT are less likely to believe that the demands imposed by a given situation exceed their available coping resources.

Aligned with our guiding definition and calls for scholars to provide evidence on the link between MT and performance, we revealed initial evidence to support the importance of MT for performance in three achievement settings. Among a sample of employees in Study 3, we showed that MT was both directly and indirectly (i.e., via a reduction in perceived distress and an increase in perceived coping ability) related to supervisor-rated performance. We extended these findings in Study 5 to objective performance in terms of success or failure in a selection test within the military context. The results of Study 5 were particularly encouraging, given that our short and direct assessment of MT emerged as the only significant predictor of sustaining high performance alongside the three hardiness and general self-efficacy (collectively, the 4Cs model of MT; Clough et al., 2002). In Study 4, we found that MT was directly related to both academic and social goal progress over a university semester among tertiary students. In these studies performance was assessed retrospectively over the past month by an informant (Study 3), prospectively over the course of a University semester (Study 4), and based on multiple and accumulating tasks over a six-week period (Study 5). Nevertheless, it is important to recognize that additional variance in performance remained unexplained and therefore the influence of MT was small but nevertheless important.

The third key contribution of this research relates to the traitness of MT. Assessments of MT have traditionally relied on a general approach in which participants were asked to reflect on their typical thoughts, emotions, or behaviors thereby emphasizing a trait perspective (cf. Clough et al., 2002). However, MT may be better conceptualized as a state-like concept such that it can both vary and have enduring properties across situations and
time (Harmison, 2011). Our research is the first to directly test this hypothesis. Using a weekly assessment timeframe with tertiary students in Study 4, we found that approximately 44% of the total variance in MT was due to between-person differences, with the remaining 56% attributable to within-person variability across the 10 weeks of the university semester. In other words, there was slightly more variability across situations than between individuals thereby supporting the hypothesis that MT may be best conceptualized as a state-like concept. These findings indicated that MT may not be as stable (i.e., trait) as previously hypothesized (Clough et al., 2002) and reported (Hardy et al., in press). From the perspective of an integrative science of personality psychology (McAdams & Pals, 2006), and specifically, a contemporary conceptualization of mental toughness which encompasses multiple personality layers (Coulter et al., 2013), these findings support an interpretation of MT as a characteristic adaptation, that is, a contextualized expression of dispositional traits that are activated or shaped by contextual or social factors (e.g., motives, values, coping styles, personal strivings, self-beliefs).

**Research Strengths and Limitations, and Conclusion**

The key strengths of this research included the use of samples from four different achievements contexts, and evidence to support different types of validity for the MT concept in terms of theoretical expectations (e.g., predict performance). Nevertheless, the key contributions of our research should be considered in light of the methodological limitations. With the exception of Studies 3 (i.e., informant-rated in role performance) and 5 (i.e., selection test outcome), our reliance on self-reported data introduced concerns associated with common method biases (Podsakoff, MacKenzie, Lee, & Podsakoff, 2003). Additionally, we relied on non-experimental designs, which do not permit causality interpretations of the relationships among study variables. Although we sampled participants from four different achievement contexts, our research was conducted solely within an Australian context using
convenience samples thus limiting any claims as to the generalizability of our findings. Future research would do well to replicate and extend our research by focusing on samples beyond the Western world to places such as the Middle East and Asia in which MT has gained interest yet relies on outdated conceptualizations and measurement tools. Finally, we did not explicitly examine MT in relation to major assaults of one’s normal functioning (e.g., serious injury or health issue, death of a loved one, failed a course). This complex aspect of our conceptualization of MT requires attention in future research.

In summary, the series of studies detailed in this paper are both timely and important because they challenge existing conceptualizations of MT as a multidimensional concept and offer a conceptual refinement based on nomological validity evidence with several diverse samples of performers. The results of Study 5, in particular, indicated that our direct assessment of MT performed better in terms of predictive validity than the indirect approach in which individual facets of the 4Cs conceptual model (Clough et al., 2002) were measured. Despite these encouraging findings, additional research is required to examine the incremental validity of MT over related concepts (see Table 2) and other umbrella concepts such as psychological capital (Luthans et al., 2007). It is not unreasonable to expect conceptual overlap between broad concepts and related specific attributes (e.g., Study 5: MT and self-efficacy). We expect this popular, umbrella concept to continue to receive scholarly attention despite such conceptual overlap, so the distinctiveness of MT must be examined in future research.
References


This article is protected by copyright. All rights reserved.


Keyes & J. Haidt (Eds.). *Flourishing: Positive psychology and the life well-lived* (pp.

motivation, controlled motivation, and goal progress. *Journal of Personality, 76*, 1201–
1229. doi: 10.1111/j.1467-6494.2008.00519.x

498. doi: 10.1207/S15327965PLB1084_01

metric and scalar invariance across mental health status and gender. *Journal of
Psychopathology and Behavioral Assessment, 34*, 48-57. doi: 10.1007/s10862-011-
9266-1


In J. Greenberg (Ed.), *Organizational behavior: The state of the science* (pp. 395-423).
Mahwah, NJ: Lawrence Erlbaum.

867-890. doi: 10.1177/0149206307307636

Questionnaire (BRSQ): Instrument development and initial validity evidence. *Journal
of Sport & Exercise Psychology, 30*, 323-355.

Sydney: The Psychology Foundation of Australia Inc.


This article is protected by copyright. All rights reserved.


This article is protected by copyright. All rights reserved.
### Table 1. Definitions of mental toughness from English language and psychology research.

<table>
<thead>
<tr>
<th>Source</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Clough, Earle, and Sewell (2002, p. 38)</td>
<td>Mentally tough individuals tend to be 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 unshakeable faith that they control their own destiny, these individuals can remain relatively unaffected by competition of adversity.</td>
</tr>
<tr>
<td>Jones, Hanton, and Connaughton (2002, p. 209)</td>
<td>Mental toughness is having the natural or developed edge that enables you to: (i) generally, cope better than your opponents with the many demands (competition, training, lifestyle) that sport places on a performer; (ii) specifically, be more consistent and better than your opponents in remaining determined, focused, confident, and in control under pressure.</td>
</tr>
<tr>
<td>Thelwell, Weston, and Greenlees (2005)</td>
<td>Mental toughness is having the natural or developed edge that enables you to: (i) always [emphasis added], cope better than your opponents with the many demands (competition, training, lifestyle) that sport places on a performer; (ii) specifically, be more consistent and better than your opponents in remaining determined, focused, confident, and in control under pressure.</td>
</tr>
<tr>
<td>Coulter, Mallett, and Gucciardi (2010, p. 715)</td>
<td>Mental toughness is the presence of some or the entire collection of experientially developed and inherent values, attitudes, emotions, cognitions, and behaviours that influence the way in which an individual approaches, responds to, and appraises both negatively and positively construed pressures, challenges, and adversities to consistently achieve his or her goals.</td>
</tr>
<tr>
<td>Clough and Strycharczyk (2012, p. 1)</td>
<td>The quality which determines in large part how people deal effectively with challenge, stressors and pressure...irrespective of prevailing circumstances.</td>
</tr>
</tbody>
</table>
| Merriam-Webster Dictionary | Mental: “of or relating to the mind.”  
Tough: “a strong or firm texture but flexible and not brittle.” |
| Oxford Dictionary | Mental: “relating to the mind.”  
Tough: “(of a substance or object) strong enough to withstand adverse conditions or rough handling” and “able to endure hardship or pain.” |
| Cambridge Dictionary | Mental: “relating to the mind, or involving the process of thinking.”  
Tough: “strong, not easily broken or made weaker or defeated.” |
<table>
<thead>
<tr>
<th>Key Dimension</th>
<th>Definition</th>
<th>Supporting Theory and Research</th>
</tr>
</thead>
<tbody>
<tr>
<td>Generalized Self-efficacy</td>
<td>A belief in your abilities to achieve success in your achievement context</td>
<td>Self-efficacy theory (Bandura, 1997)</td>
</tr>
<tr>
<td>Buoyancy</td>
<td>The ability to effectively execute to the required skills and processes in response to the challenges and pressures of everyday life.</td>
<td>Academic and workplace buoyancy (Martin &amp; Marsh, 2008, 2009)</td>
</tr>
<tr>
<td>Success Mindset</td>
<td>The desire to achieve success and ability to act upon this motive</td>
<td>Grit (Duckworth, Peterson, Matthews, &amp; Kelly, 2007); Hope theory (i.e., ‘the will’; Snyder et al., 1991)</td>
</tr>
<tr>
<td>Optimistic Style</td>
<td>The tendency to expect positive events in the future, and attribute positive causes and outcomes to different events in their lives.</td>
<td>Optimism (Scheier &amp; Carver, 1985) and explanatory style (Seligman, 1991)</td>
</tr>
<tr>
<td>Context Knowledge</td>
<td>An awareness and understanding of the performance context, and how to apply this knowledge in achieving success or reaching one’s goals</td>
<td>Cognitive theories of wisdom (e.g., Sternberg, 1998); hope theory (i.e., ‘the ways’; Snyder et al., 1991); and performance intelligence (Jones, 2012)</td>
</tr>
<tr>
<td>Emotion Regulation</td>
<td>An awareness of and ability to use emotionally relevant processes to facilitate optimal performance and goal attainment</td>
<td>Process (Gross &amp; Thompson, 2007) and goal-oriented models (e.g., Larson, 2000) of emotion regulation</td>
</tr>
<tr>
<td>Attention Regulation</td>
<td>The ability to focus on what is relevant while ignoring irrelevant information</td>
<td>Cognitive control perspectives (e.g., Robinson, Schmeichel, &amp; Inzlicht, 2010), and executive functions (e.g., Zelazo, Carter, Reznick, &amp; Frye, 1997)</td>
</tr>
</tbody>
</table>
Table 3. Summary of fit indices for measurement models examined with the calibration sample, and context specific samples for cross-validation purposes in Study 2.

<table>
<thead>
<tr>
<th>Models</th>
<th>$\chi^2$</th>
<th>df</th>
<th>p</th>
<th>CFI</th>
<th>TLI</th>
<th>SRMR</th>
<th>RMSEA (90% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Model 1: Unidimensional model (21 items)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Calibration sample ($n = 418$)</td>
<td>536.58</td>
<td>189</td>
<td>.000</td>
<td>.917</td>
<td>.908</td>
<td>.040</td>
<td>.066 (.060 to .073)</td>
</tr>
<tr>
<td>Athletes ($n = 445$)</td>
<td>609.01</td>
<td>189</td>
<td>.000</td>
<td>.901</td>
<td>.890</td>
<td>.043</td>
<td>.071 (.064 to .077)</td>
</tr>
<tr>
<td>Tertiary students ($n = 500$)</td>
<td>628.49</td>
<td>189</td>
<td>.000</td>
<td>.899</td>
<td>.888</td>
<td>.046</td>
<td>.068 (.062 to .074)</td>
</tr>
<tr>
<td>Employees ($n = 550$)</td>
<td>534.91</td>
<td>189</td>
<td>.000</td>
<td>.920</td>
<td>.911</td>
<td>.042</td>
<td>.058 (.052 to .063)</td>
</tr>
<tr>
<td><strong>Model 2: First-order model</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Calibration sample ($n = 418$)</td>
<td>347.28</td>
<td>168</td>
<td>.000</td>
<td>.951</td>
<td>.938</td>
<td>.036</td>
<td>.054 (.047 to .062)</td>
</tr>
<tr>
<td>Athletes ($n = 445$)</td>
<td>378.33</td>
<td>168</td>
<td>.000</td>
<td>.950</td>
<td>.938</td>
<td>.040</td>
<td>.053 (.046 to .060)</td>
</tr>
<tr>
<td>Tertiary students ($n = 500$)</td>
<td>411.03</td>
<td>168</td>
<td>.000</td>
<td>.944</td>
<td>.930</td>
<td>.045</td>
<td>.054 (.047 to .060)</td>
</tr>
<tr>
<td>Employees ($n = 550$)</td>
<td>362.09</td>
<td>168</td>
<td>.000</td>
<td>.955</td>
<td>.944</td>
<td>.039</td>
<td>.046 (.039 to .052)</td>
</tr>
<tr>
<td><strong>Model 3: Higher-order model</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Calibration sample ($n = 418$)</td>
<td>438.95</td>
<td>182</td>
<td>.000</td>
<td>.939</td>
<td>.929</td>
<td>.038</td>
<td>.058 (.051 to .065)</td>
</tr>
<tr>
<td>Athletes ($n = 445$)</td>
<td>434.78</td>
<td>182</td>
<td>.000</td>
<td>.941</td>
<td>.931</td>
<td>.039</td>
<td>.056 (.049 to .063)</td>
</tr>
<tr>
<td>Tertiary students ($n = 500$)</td>
<td>471.07</td>
<td>182</td>
<td>.000</td>
<td>.934</td>
<td>.923</td>
<td>.046</td>
<td>.056 (.050 to .063)</td>
</tr>
<tr>
<td>Employees ($n = 550$)</td>
<td>417.38</td>
<td>182</td>
<td>.000</td>
<td>.946</td>
<td>.937</td>
<td>.039</td>
<td>.048 (.042 to .055)</td>
</tr>
<tr>
<td><strong>Model 4: Unidimensional model (8 items)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Calibration sample ($n = 418$)</td>
<td>43.93</td>
<td>20</td>
<td>.001</td>
<td>.975</td>
<td>.965</td>
<td>.031</td>
<td>.054 (.032 to .075)</td>
</tr>
<tr>
<td>Athletes ($n = 445$)</td>
<td>39.65</td>
<td>20</td>
<td>.005</td>
<td>.980</td>
<td>.972</td>
<td>.027</td>
<td>.047 (.025 to .068)</td>
</tr>
<tr>
<td>Tertiary students ($n = 500$)</td>
<td>47.75</td>
<td>20</td>
<td>.000</td>
<td>.973</td>
<td>.963</td>
<td>.031</td>
<td>.053 (.034 to .072)</td>
</tr>
<tr>
<td>Employees ($n = 550$)</td>
<td>65.97</td>
<td>20</td>
<td>.000</td>
<td>.968</td>
<td>.955</td>
<td>.032</td>
<td>.065 (.048 to .082)</td>
</tr>
</tbody>
</table>

Note: A Model 1: unidimensional model of the 21 items; B Model 2: correlated 7-factor model of the 21 items; C Model 3: second-order model of the 21 items with a core mental toughness factor accounting for the variance between the seven lower-order factors; D Model 4: unidimensional model of the 8 items; $\chi^2$ = chi square; df = degrees of freedom; CFI = comparative fit index; IFI = incremental fit index; TLI = Tucker-Lewis index; SRMR = standardized root mean residual; RMSEA = root mean square error of approximation; # denotes not positive definite covariance matrix; residual error terms were not correlated.
<table>
<thead>
<tr>
<th></th>
<th>Calibration sample (n = 418)</th>
<th>Athletes (n = 445)</th>
<th>Tertiary students (n = 500)</th>
<th>Employees (n = 550)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>λ</td>
<td>Θ</td>
<td>λ</td>
<td>Θ</td>
</tr>
<tr>
<td><strong>Self-Belief</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>When faced with challenge or adversity, I believe in my ability to perform well</td>
<td>.97</td>
<td>.06</td>
<td>.90</td>
<td>.19</td>
</tr>
<tr>
<td>I believe in my ability to achieve my goals</td>
<td>.86</td>
<td>.25</td>
<td>.87</td>
<td>.25</td>
</tr>
<tr>
<td>I believe in my ability to consistently produce high levels of performance</td>
<td>.86</td>
<td>.27</td>
<td>.86</td>
<td>.26</td>
</tr>
<tr>
<td><strong>Attention Regulation</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>I am able to direct my attention towards relevant information</td>
<td>.99</td>
<td>.02</td>
<td>1.02</td>
<td>-.03</td>
</tr>
<tr>
<td>I remain focused on the task at hand despite adversity or challenge</td>
<td>.74</td>
<td>.45</td>
<td>.68</td>
<td>.53</td>
</tr>
<tr>
<td>I am able to regulate my focus when performing tasks</td>
<td>.83</td>
<td>.31</td>
<td>.79</td>
<td>.38</td>
</tr>
<tr>
<td><strong>Emotion Regulation</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>I am able to use my emotions to perform the way I want to</td>
<td>.93</td>
<td>.14</td>
<td>.97</td>
<td>.06</td>
</tr>
<tr>
<td>I am able to effectively manage my emotions during times of adversity</td>
<td>.71</td>
<td>.50</td>
<td>.74</td>
<td>.45</td>
</tr>
<tr>
<td>When I am performing well, my emotions do not get the better of me</td>
<td>.73</td>
<td>.46</td>
<td>.77</td>
<td>.40</td>
</tr>
<tr>
<td><strong>Success Mindset</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>I make things happen</td>
<td>.78</td>
<td>.40</td>
<td>.79</td>
<td>.38</td>
</tr>
<tr>
<td>I am determined to push myself towards higher goals</td>
<td>.81</td>
<td>.34</td>
<td>.83</td>
<td>.31</td>
</tr>
<tr>
<td>I strive for continued success</td>
<td>.86</td>
<td>.25</td>
<td>.86</td>
<td>.27</td>
</tr>
<tr>
<td><strong>Context Knowledge</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>I am aware of what I need to do to be successful</td>
<td>.90</td>
<td>.19</td>
<td>.85</td>
<td>.27</td>
</tr>
<tr>
<td>I effectively execute my knowledge of what is required to achieve my goals</td>
<td>.97</td>
<td>.06</td>
<td>.94</td>
<td>.13</td>
</tr>
<tr>
<td>I have deep knowledge of the area in which I perform</td>
<td>.63</td>
<td>.60</td>
<td>.72</td>
<td>.49</td>
</tr>
<tr>
<td><strong>Buoyancy</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>I consistently overcome adversity</td>
<td>.78</td>
<td>.40</td>
<td>.77</td>
<td>.41</td>
</tr>
<tr>
<td>I enjoy performing when things get tough</td>
<td>.81</td>
<td>.34</td>
<td>.74</td>
<td>.45</td>
</tr>
<tr>
<td>I am able to execute appropriate skills or knowledge when challenged</td>
<td>.83</td>
<td>.31</td>
<td>.80</td>
<td>.35</td>
</tr>
<tr>
<td><strong>Optimism</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>I can find a positive in most situations</td>
<td>.79</td>
<td>.38</td>
<td>.75</td>
<td>.43</td>
</tr>
<tr>
<td>I think about what can be done rather than what has happened</td>
<td>.86</td>
<td>.54</td>
<td>.64</td>
<td>.59</td>
</tr>
<tr>
<td>I can manage my thoughts to be productive</td>
<td>.65</td>
<td>.58</td>
<td>.63</td>
<td>.61</td>
</tr>
<tr>
<td>I think about what can be done rather than what has happened</td>
<td>.71</td>
<td>.49</td>
<td>.71</td>
<td>.49</td>
</tr>
</tbody>
</table>
Table 5. Standardized factor loadings (λ) and residual variances (Θ) of the 8-item unidimensional model of mental toughness in Study 2.

<table>
<thead>
<tr>
<th>Statement</th>
<th>Calibration sample (n = 418)</th>
<th>Athletes (n = 445)</th>
<th>Tertiary students (n = 500)</th>
<th>Employees (n = 550)</th>
</tr>
</thead>
<tbody>
<tr>
<td>I believe in my ability to achieve my goals</td>
<td>.75  .43</td>
<td>.68  .54</td>
<td>.70  .51</td>
<td>.73  .47</td>
</tr>
<tr>
<td>I am able to regulate my focus when performing tasks</td>
<td>.77  .40</td>
<td>.81  .35</td>
<td>.70  .51</td>
<td>.80  .36</td>
</tr>
<tr>
<td>I am able to use my emotions to perform the way I want to</td>
<td>.68  .54</td>
<td>.70  .51</td>
<td>.63  .61</td>
<td>.61  .63</td>
</tr>
<tr>
<td>I strive for continued success</td>
<td>.68  .54</td>
<td>.70  .52</td>
<td>.66  .57</td>
<td>.67  .55</td>
</tr>
<tr>
<td>I effectively execute my knowledge of what is required to achieve my goals</td>
<td>.72  .41</td>
<td>.73  .47</td>
<td>.77  .40</td>
<td>.75  .44</td>
</tr>
<tr>
<td>I consistently overcome adversity</td>
<td>.72  .48</td>
<td>.67  .56</td>
<td>.65  .58</td>
<td>.68  .53</td>
</tr>
<tr>
<td>I am able to execute appropriate skills or knowledge when challenged</td>
<td>.67  .56</td>
<td>.66  .57</td>
<td>.63  .60</td>
<td>.66  .56</td>
</tr>
<tr>
<td>I can find a positive in most situations</td>
<td>.64  .60</td>
<td>.61  .63</td>
<td>.57  .68</td>
<td>.56  .68</td>
</tr>
</tbody>
</table>
Table 6. Standardized parameter estimates for the MLSEM exploring mental toughness as a predictor of goal progress, thriving, and psychological health in Study 3b.

<table>
<thead>
<tr>
<th>Structural Path</th>
<th>Within Person (Level 1)</th>
<th>Between Person (Level 2)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>β</td>
<td>SE</td>
</tr>
<tr>
<td><strong>Direct Effects</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MT → DASS</td>
<td>-.32***</td>
<td>.05</td>
</tr>
<tr>
<td>MT → EMO</td>
<td>.58***</td>
<td>.04</td>
</tr>
<tr>
<td>MT → THRIVE</td>
<td>.64***</td>
<td>.03</td>
</tr>
<tr>
<td>MT → AcGOAL</td>
<td>.38***</td>
<td>.03</td>
</tr>
<tr>
<td>MT → ScGOAL</td>
<td>.18***</td>
<td>.04</td>
</tr>
<tr>
<td><strong>Latent Correlations</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>EMO ←→ DASS</td>
<td>-.33***</td>
<td>.05</td>
</tr>
<tr>
<td>THRIVE ←→ DASS</td>
<td>-.18***</td>
<td>.05</td>
</tr>
<tr>
<td>THRIVE ←→ EMO</td>
<td>.54**</td>
<td>.04</td>
</tr>
<tr>
<td>AcGOAL ←→ DASS</td>
<td>-.06</td>
<td>.04</td>
</tr>
<tr>
<td>AcGOAL ←→ EMO</td>
<td>.23***</td>
<td>.04</td>
</tr>
<tr>
<td>AcGOAL ←→ THRIVE</td>
<td>.18***</td>
<td>.04</td>
</tr>
<tr>
<td>ScGOAL ←→ DASS</td>
<td>-.12**</td>
<td>.04</td>
</tr>
<tr>
<td>ScGOAL ←→ EMO</td>
<td>.23**</td>
<td>.04</td>
</tr>
<tr>
<td>ScGOAL ←→ THRIVE</td>
<td>.14**</td>
<td>.04</td>
</tr>
<tr>
<td>ScGOAL ←→ AcGOAL</td>
<td>.27***</td>
<td>.03</td>
</tr>
</tbody>
</table>

Note: MT = mental toughness; THRIVE = thriving; AcGOAL = academic goal progress; ScGOAL = social goal progress; DASS = negative emotional states; EMO = positive emotions; * p < .05, ** p < .01, *** p < .001