Progress Toward Construct Validation of the Sports Mental Toughness Questionnaire (SMTQ)

Michael Sheard1, Jim Golby2, and Anna van Wersch2

1York St John University, York, UK, 2University of Teesside, Middlesbrough, UK

Abstract. This study examines the construct validity of an original self-report instrument for the assessment of mental toughness: the Sports Mental Toughness Questionnaire (SMTQ). Two independent studies supported a three-factor (Confidence, Constancy, and Control) 14-item model for the SMTQ. With a sample of 633 athletes (427 males, 206 females; M age = 21.5 years; SD = 5.48), drawn from 25 sport classifications, and competing at international, national, county and provincial, or club and regional standards, the first study utilized item development and exploratory factor analytic techniques to establish the psychometric properties of the SMTQ. Study 2 employed confirmatory factor analytic techniques with an independent sample of 509 sports performers (351 males, 158 females; M age = 20.2 years; SD = 3.35), competing at the aforementioned standards, and representative of 26 sports. Confirmatory analysis using structural equation modeling confirmed the overall structure. A single factor underlying mental toughness (Gmt) was identified with higher-order exploratory factor analysis using the Schmid-Leiman procedure. Collectively, satisfying absolute and incremental fit-index benchmarks, the inventory was shown to possess satisfactory psychometric properties, with adequate reliability, divergent validity, and discriminative power. The results revealed promising features of the SMTQ, lending preliminary support to the instrument’s factorial validity and reliability. Further construct validation of the SMTQ is recommended, including its use as an index for evaluating the effect of intervention programs.

Keywords: mental toughness, Sports Mental Toughness Questionnaire (SMTQ), principal axis factoring analysis, confirmatory factor analysis, Schmid-Leiman

Introduction

Research has identified “mental toughness” as a crucial attribute for success in competitive sport and the development of champion sport performers (e.g., Durand-Bush & Salmela, 2002; Gould, Dieffenbach, & Moffett, 2002). Yet, despite a general definitional consensus reflecting the cognitive-behavioral multivariate nature of the construct (see Crust, 2007, for a review), insufficient effort has been devoted to the development of a reliable and valid measure of mental toughness in sport. Though attention has been given to the construction of scales measuring the construct, little supporting psychometric evidence is available for them.

The Mental Toughness Questionnaire 48 (MTQ48; Clough, Earle, & Sewell, 2002) contains 48 items that are scored on a five-point Likert scale ranging from strongly agree to strongly disagree, with an average completion time of between 10 and 15 min. The scale yields scores for overall mental toughness and for each of its subscales: commitment, emotional control, life control, challenge, interpersonal confidence, and confidence in abilities. However, there is little published evidence of the scale’s psychometric properties. Its factor structure, in particular, has not been rigorously scrutinized using exploratory and confirmatory techniques. Further, the scale’s authors offered little rationale for the association with hardiness, which, ultimately, is the basis for the instrument’s subscales.

The Mental Toughness Inventory (MTI; Middleton, Marsh, Martin, Richards, & Perry, 2004) is a 65-item instrument that measures 12 components of mental toughness along with global mental toughness. The MTI appears to have been developed from a sound theoretical base and has been evaluated via a construct validation framework. However, few other details of the scale are available. Moreover, it can be considered a limitation that the scale has been validated using only elite sports high school athletes with a mean age of 14 years (range 12 to 19 years). Further testing of the MTI is needed to determine its psychometric properties.

The 42-item Psychological Performance Inventory (PPI; Loehr, 1986) was one of the first measures to include cognitive-behavioral and self-evaluation dimensions. Although Loehr offered a persuasive discussion of the scale, the conceptual and theoretical basis for the instrument was not strong and, in particular, he presented no psychometric support for its use. Indeed, psychometric analysis of the PPI (Golby, Sheard, & van Wersch, 2007) found little support for the instrument’s proposed seven-factor structure.

Golby et al. (2007) extended their psychometric work on the PPI by successfully developing an alternative instru-
ment (PPI-A). They ultimately noted four factors of mental toughness; these were labeled Determination, Self-belief, Positive cognition, and Visualization. Confirmatory factor analysis (CFA) showed good support for the four-factor PPI-A. However, a potential limitation of the scale is that it does not include a measure of control, a characteristic repeatedly identified in the mental toughness literature (see Jones, Hanton, & Connaughton, 2002). Consequently, the present research reports two studies that sought to examine the development of an original multidimensional measure of mental toughness in sport.

Study 1 focused on the initial construction of the Sports Mental Toughness Questionnaire (SMTQ), the preliminary determination of the factor structure, and the psychometric properties of the instrument. Study 2 employed confirmatory factor analytic techniques to provide further support for the generated model and the psychometric properties of the instrument.

Materials and Methods

Participants and Procedure

A total of 1,142 male and female athletes ranging from 16 to 63 years of age participated in the SMTQ scale development and evaluation phases. Two independent samples were employed. Sample 1 consisted of 633 performers, 16 to 63 years old (427 males, 206 females; M age = 21.5 years; SD = 5.48), drawn from 25 sport classifications. This sample was employed in Study 1. Sample 2 was an independent sample of 509 athletes, 18 to 48 years old (351 males, 158 females; M age = 20.2 years; SD = 3.35), drawn from 26 sport classifications. These athletes were used in Study 2. Both samples consisted of performers competing at international, national, county and provincial, or club and regional standards.

Institutional ethics approval was obtained. Volunteers were recruited via personal communication, letter, and e-mail invitations to a broad sample of clubs and governing bodies in the United Kingdom (contact first author for a sample report). Ethical procedures conforming to standards set by the British Psychological Society (2006) were adhered to throughout the research process. All participants provided informed consent and were assured confidentiality. Participants were naïve to the research hypotheses and completed the questionnaires in their respective training camps. There were six stages to the data collection:

Study 1

Stage 1: Item Development

Using raw data themes and quotes from qualitative studies (see Crust, 2007), a corpus of sport-relevant items was developed. Guidelines for item wording (Clark & Watson, 1995) were followed. An initial pool of 53 items was administered to five female and five male athletes, and to 10 coaches working in a variety of sports (contact first author for the set of 53 items). Using a dichotomous scale (applicable vs. inapplicable), coaches were instructed to assess the applicability of each item to their respective sport. Using the ratings provided by athletes and coaches, and their numerous comments, several items were rewritten in order to improve their clarity and to broaden their applicability across sports. Experts involved in previous mental toughness investigations reviewed the resulting bank of items. They retained 18 items.

Stage 2: Exploratory Factor Analysis

Responses to the SMTQ items were made on a four-point Likert scale anchored by not at all true and very true. To examine the common variance within the factor structure of the SMTQ, and to eliminate error variance associated with it (Tabachnick & Fidell, 2007), principal axis factoring analysis (PAF) was performed on the retained 18 items with SPSS v14.0 (SPSS Inc., Chicago, IL) using oblique rotation (Promax with Kaiser normalization), which allows for correlation among subscales (Costello & Osborne, 2005; Russell, 2002). To ensure minimal ambiguity between factors, criteria for an acceptable factor solution were (a) that factors have a minimum eigenvalue of 1, (b) the exclusion of pattern coefficients below .40, (c) the exclusion of items loading .40 or more if there was cross-loading greater than .30 on any other factors, and (d) that there should be a minimum of three items on each factor (Tabachnick & Fidell, 2007).

Results

PAF yielded a four-factor structure that explained 38.9% of the variance. Two items had pattern coefficients of less than .40, and were removed from subsequent analysis. Because this resulted in a factor with fewer than three items, two further items were removed from the model. PAF was applied to the retained 14 items. The scree test suggested that three factors should be extracted from the data set. Velicer’s MAP and parallel analysis tests were not conducted because of their tendency to underfactor and overfactor, respectively (Hayton, Allen, & Scarcella, 2004; Leshem & Glicksohn, 2007). In total, the three factors accounted for 40.7% of the variance (23.8, 12.0, and 4.9%, respectively) in the data set (Table 1). In the unrotated solution, all items loaded on a single factor providing evidence for a general factor of mental toughness. The SMTQ factors were labeled Confidence, Constancy, and Control. The 14-item inventory is presented in Table 2.

Pearson correlations among the three factors were all positive and statistically significant (p < .01): Confidence...
Study 2

A Theory-Based Model of Mental Toughness

The approach to mental toughness and the corresponding inventory introduced in the present research offers a dispositional perspective of the construct. Moreover, the construct validation of mental toughness undertaken in this research is in the spirit of what Seligman and Csikszentmihalyi (2000) have baptized “positive psychology”. A central tenet of the positive psychology paradigm is that stressors, adversity, and other inordinate demands are inherent to the human condition. However, the paradigm assumes that there are also sources of strength, by which this condition can be endured and even transcended. Physical, emotional, and social stressors can stimulate growth and strengthening in many individuals. Such people are often able to tap into previously unknown capacities, perspectives, and virtues. Thus, positive psychology is an obviously desirable frame of reference within which to examine the construct validation of mental toughness.

Mentally tough performers have the ability to bounce back from stressful experiences, such as competitive sport, quickly and effectively. Such an outcome is likely to be facilitated by athletes’ possession of relatively enduring characteristics; for example, optimism, hardiness, and positive affectivity are highly desirable dispositional tendencies that can predispose a situation-specific response. When confronting a challenge, individuals high in optimism, hardiness, and positive affectivity tend to approach it with confidence and persistence (Golby & Sheard, 2004;
Sheard & Golby, 2006b). Such persons feel engaged in, and feel they have influence over, whatever they are doing, deriving positive emotions from their involvement (see Forgas, 2006; Maddi, 2006). This perspective facilitates adaptive solution-focused behaviors, resulting in people concluding that the challenges facing them can be overcome.

Studies have revealed relationships, but conceptual distinctions, between mental toughness and hardiness (e.g., Golby, Sheard, & Lavallee, 2003), optimism (e.g., Nicholls, Polman, Levy, & Backhouse, 2008), and positive affectivity (Sheard & Golby, 2006a).

Materials and Methods

Stage 3: CFA

Sample 2 completed the 14-item SMTQ. CFA was conducted using LISREL 8.14 (Jöreskog & Sörbom, 1993), incorporating the maximum likelihood method (Chou & Bentler, 1995).

Stage 4: Construction of a Higher-Order Factor Model

Higher-order factor analysis takes into account the reality of correlated factors by allowing for the extraction of variance accounted for by the general factor (Thompson, 2004). The residualized domain-specific factor loading of each variable can be specified by partialing out the variable’s “g” loading and uniqueness. This transformation is achieved by the Schmid-Leiman procedure (Schmid & Leiman, 1957). It has been strongly recommended that higher-order model solutions should be subjected to a Schmid-Leiman transformation, prior to model solution interpretations (Gignac, 2007).

Stage 5: Divergent Validity

Participants also completed the Personal Views Survey III-R (PVS III-R; Maddi & Khoshaba, 2001), the revised Life Orientation Test (LOT-R; Scheier, Carver, & Bridges, 1994), and the Positive and Negative Affect Schedule (PANAS; Watson, Clark, & Tellegen, 1988). These instruments were chosen as the divergent validity criterion on the basis of their acceptable psychometric properties, and their conceptual relatedness to, but hypothesized independence and distinction from, mental toughness, mentioned previously. Cronbach’s αs showed acceptable internal consistency for each of the scales: .72, .74, .71, .72, .81, and .81, for Commitment, Control, Challenge (the three hardness subscales), optimism, positive affect, and negative affect, respectively.

Figure 1. The SMTQ model tested in this investigation, which includes the higher-order estimates displayed within the figure, as well as the corresponding calculations required for performing the Schmid-Leiman transformation of the higher-order model. Note. G_MT = General Mental Toughness; CF = Confidence; CS = Constancy; CT = Control.
Stage 6: Competitive Standard, Gender, and Age Differences

The two samples were combined to derive normative data for the SMTQ. Separate MANOVAs were used with follow-up ANOVAs to clarify the source of significant differences. When necessary, post hoc comparisons were made using the Dunn-Sidak method, which accounted for the inflated Type I error rate (Kirk, 1995). For the MANOVAs, competitive standard, gender, and age group served as the independent variables, while SMTQ subscale scores served as the dependent variables. α was set at .05.

Results

Schmid-Leiman Analyses

The higher-order General Mental Toughness (Gmt) factor accounted for 12% to 38% of the item variance. Most items loaded primarily on the higher-order Gmt factor. The higher-order factor-structure coefficients associated with the Gmt factor were .69, .71, and .75, for Confidence, Constancy, and Control, respectively, indicating that the general factor was strong. The unique latent variable regression paths associated with the first-order latent variables were .72, .71, and .66, for Confidence, Constancy, and Control, respectively. The unique latent variable variances associated with the first-order latent variables were .52, .50, and .44, for Confidence, Constancy, and Control, respectively. Thus, evidence suggests the plausibility of both a Gmt factor, and three unique first-order factors, corresponding to Confidence, Constancy, and Control. The Schmid-Leiman transformed pattern coefficients were calculated by multiplying the first-order pattern coefficients by their respective unique regression paths associated with the unique first-order latent variable (see Figure 1 for calculations).

CFA

Descriptive analyses determined that the items tended to be associated with relatively nonkurtotic distributions (i.e., < |1.0|) and no particularly conspicuous outliers were identified. Thus, the data were considered appropriately distributed for the purposes of maximum likelihood estimation. The intercorrelations, means, and standard deviations for the 14 SMTQ items completed by Sample 2 are available from the first author. The oblique three-factor model yielded absolute fit indices that indicated a good model fit: χ² (74, N = 509) = 182.56, p < .01, χ²/df = 2.47, goodness-of-fit index (GFI) = .95, adjusted goodness-of-fit index (AGFI) = .93, root-mean-square-error of approximation RMSEA = .05, root-mean-square residual (RMR) = .05. Similarly, the incremental fit indices indicated good support for the model: TLI = .91, CFI = .92, IFI = .93. The equivalent higher-order model, with three first-order factors and one second-order general factor yielded the same degree of model fit as the preceding oblique three-factor model (as expected). The three factors’ intercorrelations were all positive and statistically significant (p < .01): Confidence · Constancy = .31; Confidence · Control = .28; and Constancy · Control = .32. Each also demonstrated acceptable internal consistency: Confidence = .79, Constancy = .76, and Control = .72.

Divergent Validity

Divergent validity was demonstrated by the observed pattern of weak to moderate correlations between the SMTQ

Table 3. SMTQ means and standard deviations

<table>
<thead>
<tr>
<th>Factor</th>
<th>Confidence</th>
<th>Constancy</th>
<th>Control</th>
<th>Total mental toughness</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>M</td>
<td>SD</td>
<td>M</td>
<td>SD</td>
</tr>
<tr>
<td>Competitive standard</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>International (n = 79)</td>
<td>18.22</td>
<td>2.82</td>
<td>13.97</td>
<td>1.95</td>
</tr>
<tr>
<td>National (n = 150)</td>
<td>17.49</td>
<td>3.70</td>
<td>13.79</td>
<td>1.95</td>
</tr>
<tr>
<td>County/provincial (n = 479)</td>
<td>16.48</td>
<td>3.17</td>
<td>12.73</td>
<td>2.15</td>
</tr>
<tr>
<td>Club/regional (n = 434)</td>
<td>15.27</td>
<td>3.00</td>
<td>12.69</td>
<td>2.18</td>
</tr>
<tr>
<td>Gender</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male (n = 778)</td>
<td>17.03</td>
<td>3.12</td>
<td>12.92</td>
<td>2.27</td>
</tr>
<tr>
<td>Female (n = 364)</td>
<td>14.65</td>
<td>3.05</td>
<td>12.99</td>
<td>1.95</td>
</tr>
<tr>
<td>Age group</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>16–18 years (n = 272)</td>
<td>15.81</td>
<td>3.56</td>
<td>12.90</td>
<td>2.23</td>
</tr>
<tr>
<td>19–20 years (n = 505)</td>
<td>16.06</td>
<td>3.16</td>
<td>12.85</td>
<td>2.06</td>
</tr>
<tr>
<td>21–24 years (n = 231)</td>
<td>16.59</td>
<td>3.00</td>
<td>12.87</td>
<td>2.20</td>
</tr>
<tr>
<td>25+ years (n = 134)</td>
<td>17.50</td>
<td>3.36</td>
<td>13.49</td>
<td>2.35</td>
</tr>
</tbody>
</table>

Note. Data from combined samples (N = 1142).
and PVS III-R \((r \text{ range} = .14–.33)\), LOT-R \((r \text{ range} = .23–.38)\), and PANAS \((r \text{ range} = .12–.49)\) subscales.

**Competitive Standard, Gender, and Age Differences**

Means and standard deviations of all variables are presented in Table 3. Significant MANOVA main effects were observed for competitive standard, Wilks’ \(\lambda = .88, F(9, 2765) = 15.92, p < .001, \text{partial } \eta^2 = .04\); gender, Wilks’ \(\lambda = .88, F(3, 1138) = 53.99, p < .001, \text{partial } \eta^2 = .13\); and age group, Wilks’ \(\lambda = .96, F(9, 2765) = 4.75, p < .001, \text{partial } \eta^2 = .012\). Follow-up ANOVAs revealed significant differences in Confidence and Constancy relative to standard. County/provincial performers scored significantly higher in Confidence than their club/regional counterparts. However, these groups scored significantly lower than either international or national athletes in both Confidence and Constancy (all \(p < .001\)). Male athletes scored significantly higher than their female counterparts in both Confidence (\(p < .001\)) and Control (\(p = .03\)). Performers in the oldest age group (i.e., \(\geq 25\) years) scored significantly higher in all SMTQ factors compared to their younger counterparts (all \(p < .05\)).

Because total SMTQ scores are a composite of the scale’s three factors, separate univariate ANOVA were conducted for total mental toughness. The independent variables were the same as those used in the MANOVA. International and national performers scored significantly higher than their subelite counterparts. Male athletes scored significantly higher than female performers, and athletes in the oldest age group were characterized by significantly higher levels of total mental toughness than younger performers (all \(p < .05\)).

**Discussion**

In this study we undertook a comprehensive analysis of the construct validity of the SMTQ. A first aim of this research was to develop a questionnaire for the assessment of athletes’ mental toughness, while providing preliminary evidence for its factorial validity. Eighteen items were retained from the 53 initially created. PAF identified three factors, using criteria mentioned previously; these were labeled Confidence, Constancy, and Control. The final SMTQ consisted of 14 items. Model fit was assessed using CFA, and both absolute and incremental fit indices showed good support for the correlated three-factor SMTQ model. Higher-order factor analysis suggested evidence of both a Gmt factor and the three previously mentioned first-order factors. Divergent validity was established by a pattern of correlations between the SMTQ and positive psychological measures. Finally, the scale’s discriminative power was established. Collectively, the results offer preliminary support for the validity and reliability of the SMTQ as a measure of mental toughness.

The SMTQ has good content validity. A panel of experts assessed an initial item pool for comprehensibility by athletes. The factors that emerged subsequently from the exploratory analysis correspond meaningfully with the definitions of mental toughness within the extant literature: Confidence (vs. self-doubt), Constancy (vs. irresolute), and Control (vs. agitation) are themes encountered frequently in the recent qualitative studies undertaken in the sport domain (see Crust, 2007). Moreover, the SMTQ is the only psychometrically acceptable mental toughness instrument that includes a measure of emotional and negative energy control, a characteristic routinely identified in the mental toughness literature (see Clough et al., 2002; Jones et al., 2002; Loehr, 1986). This omission was a limitation of the PPI-A identified by Golby et al. (2007). Further, the three factors extracted in the exploratory analysis exhibited good internal consistency with each independent sample used in this study.

First-order and higher-order CFA suggested adequate model fit. The correlated three-factor model identified the unique factorial validity associated with the scores derived from the items within each of the factors. The Schmid-Leiman procedure allowed for the removal of the variance accounted for by the higher-order factor. Once the variance accounted for by Gmt was partialled out, the items that remained were assigned to theoretically consistent factors for the three-factor solution. There is something common to all the measures of mental toughness, irrespective of the factor they are intended to tap. It is important to note that the Schmid-Leiman transformation does not alter the explanatory power of the original CFA solution; for example, the proportion of unexplained variance in each item is identical before and after the transformation (Brown, 2006). Rather, the Schmid-Leiman procedure is a method of calculating the contribution of lower- and higher-order factors to the prediction of observed measures, and explanatory preference is afforded to the higher-order factors.

As emphasized by, for example, Seligman and Csikszentmihalyi (2000), the study of optimal functioning, human strengths, and positive psychological outcomes is growing rapidly. Consistent with this emphasis, mental toughness focuses considerable attention on what it takes to be an outstanding athlete and how to facilitate performance. Correlational findings from this research indicate that although mental toughness and other positive psychological constructs may share the same conceptual space, they also show sufficient and distinct independence.

Finally, subgroup differences relative to competitive standard, gender, and age indicated some variation in mental toughness levels, revealing that the SMTQ possesses good discriminative power. The implication of these findings is that, first, and subject to further validation, different sets of norms may need to be developed. Second, in addition to the many attributes that are requisite to outstanding sport performance, a psychological profile that includes...
high levels of mental toughness appeared to distinguish older, male, performers operating at the highest competitive standard (i.e., international).

Strengths of the present research were its recruitment of athletes from a variety of sports, and that studies of multidimensional questionnaires used in the sport domain (see Durand-Bush, Salmela, & Green-Demers, 2001; Gaudreau & Blondin, 2002) have reported psychometric data based on the responses of participants with a comparable age range. This research offers support for the SMTQ as a valid measure of mental toughness. The instrument possesses encouraging psychometric integrity. As construct validation is an ongoing process (Anastasi & Urbina, 1997), meaningful comparisons should now be made on SMTQ data collected over time, with the recommendation that practitioners use changes in SMTQ scores as an index for evaluating the impact of psychological skills training. This would fulfill the ultimate construct validation criterion of application in research and practice (Marsh, 2002).

Acknowledgments

The authors express their gratitude to Dr. Catherine Palmer and Dr. Ineke Vergeer, Durham University, who facilitated collection of some of the data. Parts of this study were presented during a poster session at the 2007 10th European Congress of Psychology in Prague, Czech Republic. The authors also offer sincere thanks to the anonymous reviewers who provided very helpful comments on an earlier draft of this article.

References


---

Michael Sheard

Faculty of Health & Life Sciences
York St John University
York YO31 7EX
UK
E-mail m.sheard@yorksj.ac.uk