

Further Validation of the Inventory of Mental Toughness Factors in Sport (IMTF-S)

Mark A. Stonkus¹ & Kenneth D. Royal²

¹ Clutch Athlete Training Center, Norwell, MA, USA

² Department of Clinical Sciences, College of Veterinary Medicine, North Carolina State University, Raleigh, NC, USA

Correspondence: Kenneth D. Royal, Department of Clinical Sciences, College of Veterinary Medicine, North Carolina State University, 1060 William Moore Dr., Raleigh, North Carolina, 27607, USA. Tel: 1-919-613-5100. E-mail: kdroyal2@ncsu.edu

Received: April 20, 2015 Accepted: May 27, 2015 Online Published: August 12, 2015

doi:10.5539/ijps.v7n3p35

URL: <http://dx.doi.org/10.5539/ijps.v7n3p35>

Abstract

The purpose of this study was to provide further validation a new measure of mental toughness in sport. The Inventory of Mental Toughness Factors in Sport (IMTF-S) was originally developed and validated using principal component analysis. For the present study, the psychometric properties of the IMTF-S were again evaluated, but by way of the Rasch Rating Scale Model using the same sample (n=329) of athletes, and the same 42-item instrument measuring mental toughness on a 5-point Likert-type scale (always-never). Results indicate the IMTF-S is a psychometrically-sound instrument capable of producing valid and reproducible measures of mental toughness in sports.

Keywords: mental toughness, sports, psychometrics, measurement, instrument validation

1. Introduction

1.1 Need for Sound Instrumentation

Mental toughness has been acknowledged as both a decisive factor in sport performance, as well as something that players and coaches value as being important to develop within their programs (Clough, Earle, & Sewell, 2002). It has also been determined to assist athletes in obtaining success by optimizing practice, overcoming failures, and developing the mental skills necessary to win (Norris, 1999). A past investigation produced similar findings in claiming that mental toughness, though very popular in terms of language, is actually one of the least understood concepts in the field of sport psychology (Jones, Hanton, & Connaughton, 2002).

The concept of mental toughness has been identified as one's ability to withstand adversity, pressure and stress (Clough et al., 2002; Goldberg, 1998; Jones et al., 2002; Loehr, 1995; Middleton, Marsh, Martin, Richards, & Perry, 2004a; Williams, 1988). In an attempt to define and understand mental toughness in the sport of soccer, Thelwell, Weston and Greenless (2005) found mental toughness as having the "natural or developed psychological edge that enables [one] to ... always cope better than your opponents with the many demands (competition, training, lifestyle) that soccer places on the performer" and to "be more consistent and better than your opponents in remaining determined, focused, confident, and in control under pressure" (p. 328). Goldberg (2005) offers a similar definition of mental toughness which includes factors such as an athlete's ability to cope with pressure/stress/adversity, to rebound after failure, to persist, and to be emotionally resilient. Gould, Dieffenbach and Moffet (2002), in conducting a qualitative study of 10 U.S. Olympic champions, found that mental toughness was rated as a very important component of their success (mentioned by 73% of the subjects), but stated that there was no precise definition of the term used in the study.

In providing a review of mental toughness in sport, Crust (2007) noted both the importance of the concept, and the lack of quality research and measurement of it. One of the first tools used to assess mental toughness was the Psychological Performance Inventory (PPI; Loehr, 1986). Loehr's PPI consists of 42 items and measures his proposed seven characteristics of mental toughness: self-confidence, negative energy, attention control, visual and imagery control, motivation, positive energy, and attitude control. This measurement, although conceptually logical and understandable as a measure of an athlete's mental toughness, has been criticized for lacking

psychometric properties and theoretical frameworks for its development (Crust, 2007; Mack & Ragan, 2008; Middleton, Marsh, Martin, Richards, & Perry, 2004b; Middleton et al., 2004; Sheard, Golby, & van Wersch, 2009).

Stonkus (2011) conducted a comprehensive study of the theoretical dimensions of mental toughness, and utilized the theoretical concepts to develop a new measurement of mental toughness. The investigation led to his development of the Inventory of Mental Toughness Factors in Sport (IMTF-S), which held an overall Cronbach's Alpha reliability measure of .925 (Stonkus, 2011). The purpose of the current study was to further assess the psychometric validity of the IMTF-S using the same sample but a different method of statistical analysis.

1.2 Inventory of Mental Toughness Factors in Sport (IMTF-S)

In his original study, Stonkus (2011) conducted a thorough review of the research literature related to mental toughness, and discovered six theoretical concepts from which to develop the new scale of measurement. The items generated for the Inventory of Mental Toughness Factors in Sport (IMTF-S) were comprised of themes relating to the theoretical sources of Coping (Lazarus & Folkman, 1984), Hardiness (Kobasa, 1979), Optimism (Seligman, 2006), Mindset (Dweck, 2009), Resilience (Garmezy, 1993), and Self-Efficacy (Bandura, 1977, 1995). These theoretical concepts have been shown to be intricately related to mental toughness, and it was argued that they provide the best means to operationalize and measure this construct (Stonkus, 2011).

1.3 Purpose of the Present Study

The present research sought to accomplish the primary goal of providing further evaluation and confirmation of the psychometric properties of the IMTF-S to determine how well the instrument functions relative to a diverse body of athletes. Specifically, initial efforts to validate the IMTF-S utilized traditional statistical methods, namely factor analysis. The present study utilizes a form of item response theory (IRT) measurement modeling to evaluate the psychometric properties of the IMTF-S, as this approach is more robust and offers greater insights about the quality of an instrument and its functioning relative to an intended sample. The secondary aim was to discuss the potential use of the newly developed instrument and encourage others researchers and practitioners to consider its use.

2. Methods

2.1 IMTF-S Subscales

Initially, factor analyses were performed to determine relationships among the items (Stonkus, 2011). Four primary factors emerged, which resulted in the created of four subscales on the IMTF-S. First, *Motivation*, was defined as "An athlete's ability to utilize the most functional mindset in optimizing performance". These items mostly incorporate concepts of mindset, but also hardiness, coping, and optimism, in categorizing an athlete's ability to manage thoughts and cognition associated with high-performance. Second, *Identification*, was defined as "An athlete's ability to utilize confidence and optimism in developing a strong identity to enhance performance in competition and training". The items combine the resources on optimism, self-efficacy, coping, and hardiness in categorizing the notion of an athlete's ability to manage self-image and personal expectations of him/herself in competitive situations. Next, *Negation*, was identified as "An athlete's ability to negate the dysfunctional and counterproductive influence of distractions during competition". These items incorporate the resources on resilience, coping, mindset, and hardiness in categorizing the mental toughness required for an athlete to concentrate on his/her performance in managing thoughts, emotions, and behaviors in the face of adversity (weather conditions, high-stake situations, crowd size/support, "crunch-time", etc.). Finally, *Determination*, was defined as "An athlete's ability to remain determined to work hard regardless of the situation or conditions". These items actually include all of the concepts of the sources used in generating the scale (optimism, coping, resilience, mindset, hardiness, and self-efficacy). The key to this grouping is that the items are categorized around the athlete's ability to manage effort, and to avoid giving up under stress. A sample of the items appearing on the IMTF-S are presented in Table 1. A breakdown of items comprising each subscale is presented in Table 2.

Table 1. Examples of items appearing on the IMTF-S

<i>Motivation</i>	
1) Setbacks and failure allow me to learn.	
6) Hard work involving planning will ensure my athletic success.	
<i>Identification</i>	
7) In pressure situations, my ability to execute skills increases.	
8) I have unique competitive strengths that few opponents have.	
<i>Negation</i>	
2) I am afraid of choking under pressure.	
3) I get frustrated when I make mistakes in competition.	
<i>Determination</i>	
4) I can quickly regain my composure if I have briefly lost it.	
12) When I get tired, I am not able to play my best.	

Table 2. Subscale information

Subscale	Item numbers
Motivation (n = 10)	1, 6, 11, 15, 16, 17, 25, 30, 31, 38
Identification (n = 10)	7, 8, 13, 18, 23, 24, 26, 27, 28, 29
Negation (n = 10)	2, 3, 5, 9, 10, 19, 21, 35, 36, 37
Determination (n = 10)	4, 12, 14, 20, 22, 32, 33, 34, 39, 40

2.2 Participants

A total of 329 participants were recruited from various sports teams in Massachusetts. Ages ranged from 12-34, with a mean of 17.489, (SD = 3.17). Levels included middle school athletes (ages 12-14), high school athletes (ages 14-18) collegiate athletes (ages 18-22) and amateur & professional athletes who are no longer in a school setting (ages 22-34). Of the 329 subjects, 170 (51.7%) were male, and 159 (48.3%) were female. School grade distribution was as follows: Grades 6-9 = 98, grades 10-12 = 40, college freshmen = 59, college sophomore = 55, college junior = 43, college senior = 31, post college athlete = 3. Athletic participation was comprised of the following sports: soccer, ice hockey, football, basketball, baseball, track and field, crew/rowing, lacrosse, sailing, golf, cross-country, snowboarding, field hockey, tennis, rugby, swimming/diving, skiing, wrestling, volleyball, cheerleading, and equestrian/riding. Of the 329 subjects, 99 (30%) reported being actively competitive in multiple (2 or more) sports annually, while the remaining 230 (70%) reported participation in one sport. It should be noted that only 28 (14.7%) of all post-high school subjects reported being a multi-sport athlete. No other demographic data were obtained.

Subjects were chosen based on availability and willingness to participate, and not to fill a certain quota of sports experience or participation; however these demographics reveal a wide variety of sports levels and affiliations, which should enhance the external validity of the results of the study. Additionally, because the data analysis technique employed in this study is invariant, the particulars of the sample are not important. What is important, however, is that a single unidimensional construct for measurement is discernible. This will be discussed later in this study.

2.3 Analysis

Data were analyzed with the Rasch Rating Scale Model (RRSM & Andrich, 1978). The RRSM is particularly useful because it possesses the property of invariance. Whereas traditional methods for survey validation studies are sample dependent, the RRSM can objectively measure both the latent trait and the difficulty to endorse each item without regard to the particulars of the sample (Bond & Fox, 2007; Wright & Stone, 1999). Further, the RRSM makes it possible to place both person and item measures onto the same linear continuum to discern the relationship between these two facets.

According to the RRSM model, the probability of a person n responding in category x to item i , is given by:

$$P_{xni} = \frac{\exp \sum_{j=0}^x [\beta_n - (\delta_i + \tau_j)]}{\sum_{k=0}^m \exp \sum_{j=0}^k [\beta_n - (\delta_i + \tau_j)]} \quad x = 0, 1, \dots, m \quad (1)$$

where $\tau_0 = 0$ so that $\exp \sum_{j=0}^0 [\beta_n - (\delta_i + \tau_j)] = 1$. β_n is the person's position on the variable, δ_i is the scale

value (difficulty to endorse) estimated for each item i and $\tau_1, \tau_2, \dots, \tau_m$ are the m response thresholds estimated for the $m + 1$ rating categories. Winsteps measurement software (Linacre, 2015) was used to perform the data analysis. Parameters were estimated using joint maximum likelihood estimation procedures (Wright & Masters, 1982).

3. Results

The evaluation of the psychometric properties of the IMTF-S focused on six criteria: dimensionality, reliability, rating scale quality, person measure quality, item quality, and the item hierarchy. The findings of the psychometric evaluation are presented below.

3.1 Dimensionality

A Rasch-based Principal Components Analysis (PCA) of standardized residual correlations was performed to assess dimensionality. A total of 41.8% of the variance was explained by the measures, with 30.2% of the item variance explained. The largest secondary dimension had an eigenvalue of 3.7, indicating a strength of about four items, and accounted for 5.4% of the variance. The ratio of the overall variance explained by the items relative to the largest secondary dimension was about 6:1. Collectively, this information suggests the primary dimension was both sufficiently strong in magnitude to be detectable and sufficiently independent from other possible dimensions that a primarily unidimensional measurement system could be constructed.

3.2 Reliability

Reliability and separation measures estimate the extent to which measures are reproducible under similar conditions that include similar samples. Separation measures provide a ratio for sample deviation, corrected for error, to the average estimation error (Linacre, 2011). Reliability estimates for the full instrument are high (greater than .92). Subscales provided moderate-high estimates of reliability with the least reproducible subscale reporting an estimate of .75. Table 3 provides a complete breakdown of reliability and separate measures for the overall instrument and each subscale.

Table 3. Reliability and separation measures for subscales

Subscale	Category	Reliability	Separation
Overall	Persons	.92	3.51
	Items	.99	12.04
Motivation	Persons	.75	1.73
	Items	.98	7.29
Identification	Persons	.86	2.48
	Items	.96	5.06

Negation	Persons	.84	2.32
	Items	.99	11.09
Determination	Persons	.79	1.92
	Items	.99	12.84

3.3 Rating Scale Effectiveness

The quality of a rating scale can be determined by the extent to which response options were appropriate, the categories functioned as intended, and the consistency of interpretation of items by participants (Linacre, 2002). Table 4 provides a summary of rating scale diagnostics. Results indicate the rating scale is functioning properly. Structure calibration and category measures indicate raters were able to adequately discern the ordinal nature of the rating scale and respond in a consistent manner. For the most part, respondents were able to make full use of the scale. The only exception, as indicated by slightly noise fit statistics, was the “Never” category was rarely used. Although the inclusion of the category does not result in unproductive measurement, it does not necessarily help either. Future administrations of the IMTF-S might consider collapsing the “Never” and “Rarely” categories.

Table 4. Summary of rating scale effectiveness

Category	Count	Percent	INFIT MNSQ	OUTFIT MNSQ	Structure Calibration	Category Measure
Never	360	2%	1.44	1.64	NONE	-3.12
Rarely	1,418	9%	.95	.97	-1.79	-1.45
Sometimes	4,261	28%	.95	.96	-.87	-.07
Frequently	5,509	36%	.92	.90	.64	1.43
Always	3,611	24%	.97	.97	2.03	3.28

3.4 Person Measure Quality

Person measure quality was evaluated by examining the stability of measures, size of standard errors, and fit statistics. Overall, data fit the model very well with both INFIT and OUTFIT mean square values of 1.01. The average measure was 1.06 (SD = .74) with a standard error of .21. No data were removed in order to improve data to model fit.

3.5 Item Measure Quality

Item measure quality can be determined by examining item measures, size of associated standard errors, and fit statistics. With any Rasch analysis that does not incorporate item anchoring the default item measure is .00 (SD = .82). Here, the overall standard error for items was .07, with nearly ideal fit statistics of 1.01 for both INFIT and OUTFIT mean square values. With regard to individual items, Wright and Linacre (1994) suggested fit statistics within the range of .6 and 1.4 are ideal. Each of the items appearing on the IMTF-S have fit values that accord to these ideal values. Point measure correlations examine the extent to which items adequately discriminate the amount of the latent trait under investigation. Under the Rasch measurement framework, point measure correlations should be positive to indicate adequate discriminatory abilities (Linacre, 2015). For the IMTF-S, point measure correlations ranged from .21 to .66, with a mean point measure correlation of .47 (SD = .09) for all items. Local dependency was investigated by examining the residual item correlations. Values greater than 0.3 are considered potentially dependent (Smith, 2000). A review of the residual correlation matrix revealed all items on the IMTF-S resulted in residual correlations less than 0.3, providing evidence that each item functions independently and responses to one item should not lead to a particular response on another item. A full breakdown of item measures, standard errors and fit statistics for each item is presented in Table 5.

Table 5. Item statistics

Item	Measure	SE	INFIT MNSQ	OUTFIT MNSQ	Pt. Measure Correlation
1	-.68	.07	1.23	1.31	0.25
2	.76	.06	1.05	1.05	0.50
3	2.01	.06	.95	.94	0.47
4	-.08	.07	.72	.75	0.45
5	-.02	.07	.86	.85	0.55
6	-.58	.07	1.18	1.19	0.31
7	.22	.06	.81	.88	0.42
8	.35	.06	.92	.92	0.48
9	.91	.06	1.29	1.27	0.46
10	.17	.06	1.03	1.02	0.60
11	-1.78	.09	1.12	1.06	0.32
12	1.27	.06	.96	.98	0.49
13	.10	.06	.82	.82	0.66
14	-.92	.07	1.17	1.13	0.47
15	-.17	.07	.92	1.02	0.37
16	-.26	.07	.87	1.01	0.54
17	-1.10	.08	1.32	1.21	0.50
18	-.29	.07	.82	.84	0.52
19	.13	.06	.75	.76	0.55
20	.04	.07	1.01	1.02	0.42
21	1.10	.06	1.21	1.21	0.46
22	.01	.07	1.19	1.18	0.46
23	-.23	.07	.76	.75	0.59
24	-.18	.07	.81	.80	0.56
25	-1.28	.08	.89	.88	0.37
26	.33	.06	.76	.77	0.55
27	-.50	.07	1.28	1.24	0.51
28	-.32	.07	.69	.68	0.60
29	-.23	.07	1.15	1.11	0.48
30	-1.01	.08	1.13	1.10	0.43
31	-.04	.07	1.35	1.43	0.21
32	.18	.06	.60	.60	0.57
33	-1.06	.08	1.32	1.24	0.43
34	-1.13	.08	1.11	1.10	0.43
35	.42	.06	1.21	1.21	0.44
36	1.53	.06	1.16	1.17	0.40
37	1.47	.06	1.02	1.02	0.53
38	-.78	.07	1.20	1.19	0.44
39	.96	.06	.93	.93	0.44
40	.67	.06	.85	.86	0.52

3.6 Item Hierarchy

The item map presented in Figure 1 illustrates the construct hierarchy for mental toughness in sports. Item Q11, *Passion and effort are key to achieving high performance*, was rated the easiest item to endorse among the collective sample of participants. Item Q3, *I get frustrated when I make mistakes in competition*, was rated the most difficult item to endorse among the collective sample of participants. In psychological measurement, the manner in which items arrange themselves into a hierarchical order describes the construct. In theory, items appearing at the bottom of the map should be items that most any mentally tough individual would endorse. The probability that an item would be endorsed by any survey respondent decreases as one advances up the hierarchy. Likewise, persons appearing at the top of the map indicate individuals who found it easiest to endorse the items, whereas persons appearing at the bottom of the map had a more difficult time endorsing the items. It should be noted that the person measures presented here do not necessarily indicate one’s mental toughness, but rather one’s propensity to endorse items that encapsulate the construct of mental toughness.

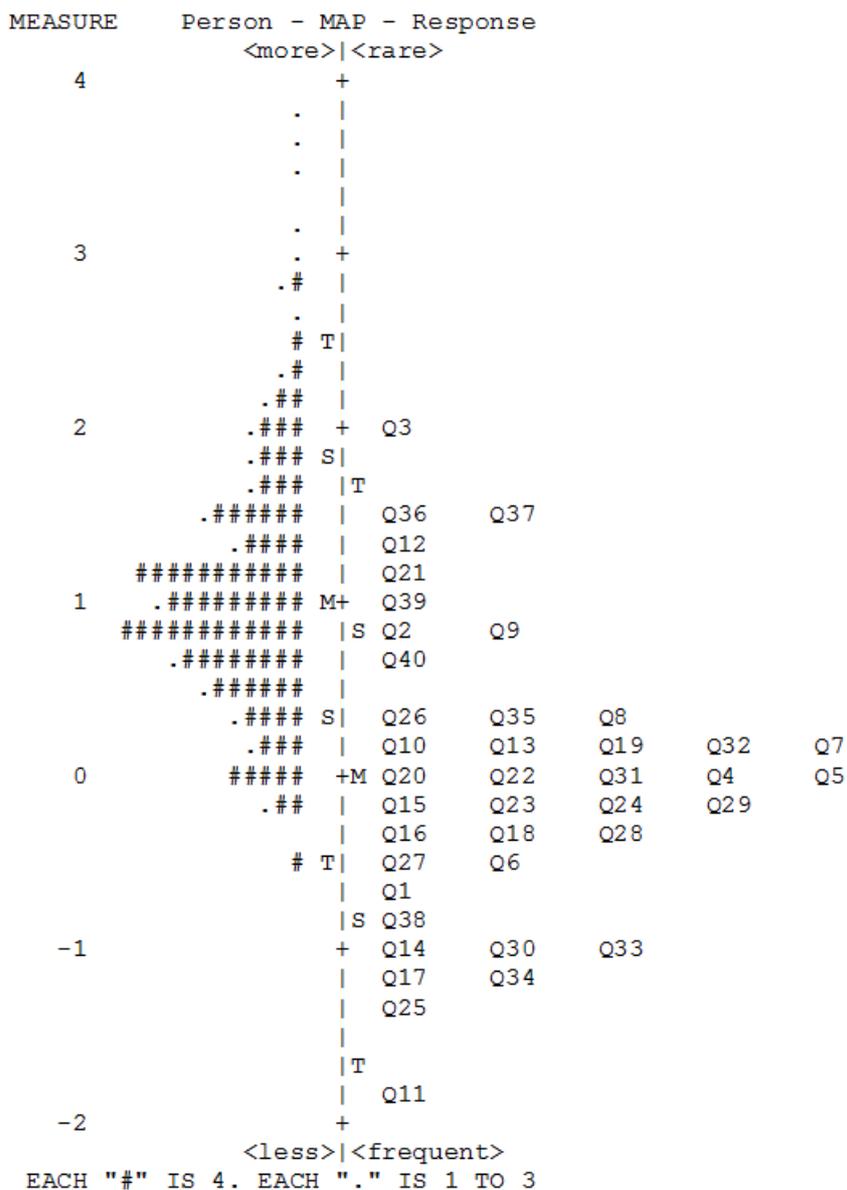


Figure 1. Construct Hierarchy of the items in the IMTF-S.

4. Discussion

4.1 Evidence of Construct Validity

The psychometric properties of the IMTF-S were evaluated with a powerful item response theory measurement model, namely the Rasch Rating Scale Model. Results for various aspects of psychometric functioning were reported. Here, we will use Messick's (1989) framework for construct validity to evaluate the findings. Messick contends that construct validity is the examination and integration of any evidence which may influence the interpretation or meaning of a score. His framework includes six components of construct validity: substantive, structural, content, generalizability, external, and consequential.

A principal components analysis of standardized residual correlations determined the Rasch dimension was both sufficient in magnitude and detection to be discernible as the primary dimension, thus meeting the requirement for unidimensionality. This finding coupled with excellent overall data to model fit statistics provides evidence that speaks to the substantive aspect of validity. An evaluation of the rating scale yielded results that indicated it was functioning as intended. This speaks to the structural aspect of validity. Item statistics were all within acceptable ranges for fit, thus speaking to the content aspect of validity. Reliability and separation estimates were quite high and provided support to the generalizability component of validity. In a very limited way external validity was supported by way of developing the instrument from theory, however, no explicit evidence to support external validity is presented. Likewise, because the study does not involve test scores or a series of consequences as a result of interpreting the measures, no evidence of the consequential aspect of validity is provided. Additionally, whereas previous research has focused primarily on attempts to measure mental toughness among athletes in a single sport, this study attempted to investigate mental toughness across a wide of sports. As such, we present no formal evidence of systematic validity, although we suspect the construct might remain intact across comparative sports. In any instances, future research is needed to investigate the stability of this construct.

4.2 Comparing Empirical Results to Theoretical Expectations

The purpose of this study was to add to the existing knowledge base and assessment of mental toughness by utilizing the most current literature and research on the topic in developing a valid instrument to measure the new conceptual model. An extensive literature review was followed by a quantitative methodology of item analysis, principal components analysis, reliability analysis, and validity measures to develop the Inventory of Mental Toughness Factors in Sport (IMTF-S) and to validate its ability to measure mental toughness in athletes. This study concludes with the conceptual and statistical evidence that the purpose of the research was successfully and adequately met.

The four-factor structure of the IMTF-S is the most recent attempt to capture the essence of mental toughness in sport. The four factors represent the most current conceptualization of mental toughness, as derived from literature, and offer perhaps the most thorough and inclusive representation of mental toughness to date. Each factor represents part or all of the theoretical foundations (Coping, Resilience, Mindset, Hardiness, Optimism, and Self-Efficacy) used in this study.

4.3 Limitations

Limitations of this research are both global and local in scope. With regard to global limitations, there is the issue of generalizability. Although the Rasch Rating Scale Model produces sample-free calibrations to define the construct of mental toughness in sport, results produced from this sample and instrument interaction cannot be generalized to specific athletes, participants of certain sports, or cultures. A more robust and representative sample is necessary to better understand what baseline measures might look like on the IMTF-S. With regard to local limitations, there are two issues in particular that might threaten validity. First, the IMTF-S was administered outside of a competitive environment, namely as they were in either a training facility or a school class/locker room. As a result, they may have been in a different state of mind than they might have been in the midst of adverse condition in sport (e.g., losing, being demoted, engaged in sudden death/overtime, standing at the free-throw line, etc.). Second, the data were self-reported from a sample of competitive athletes. Participants in this study are competitive people by nature, thus there may be a potential for some participants to try to "out score" others on the instrument. Any attempt at intentionally altering responses to inflate a score would introduce error into the measurement system. Future administrations of the IMTF-S might include a few "honesty check" questions to discern participant honesty.

4.4 Implications

The IMTF-S poses a number of potential implications for both athletes and coaches. The IMTF-S provides athletes with an objective snapshot of their mental toughness at any given point in time. The hope is that such an evaluation will assist the athlete in both understanding the value of mental toughness and the benefit of having the data to describe their current level of it. From this, the athlete can choose to train to improve his/her mental toughness by focusing on skills that would help to improve it. The benefit of the four-factor structure of the IMTF-S is that it provides a set of sub-scales which can be utilized separately in training those individual aspects of performance. The value of the IMTF-S to athletes could prove to be immeasurable, as it promotes the identification and development of an aspect of performance that is becoming more respected and valuable than physical development (Gould et al., 2002). The ability of an athlete to perform under intense episodes of stress, anxiety, frustration, and fatigue may be the key which separates good from great and novice from elite.

The importance of the IMTF-S is perhaps equally important to coaches as it is to athletes. Athletes tend to be “good soldiers” and follow the structured practice and training regimens required of them by their coaches. As such, coaches would greatly benefit from a better understanding the concept of mental toughness, their teams existing level of it, and the ways that it can be improved. The coach fosters the players’ environment; which can be either a good or a bad thing in optimizing mental toughness. If coaches are aware of the environment that they are trying to create in optimizing the performance of their players and team, then it is assumed that they would welcome the opportunity to better develop mental toughness on a daily basis. Just as athletes would benefit from more education and training on mental toughness, so would coaches. They should become more aware of what it actually is, its role in athletic performance, and how to use it to their advantage. What they must resist doing is using the IMTF-S as a screening instrument for team selection, as mental toughness should not be mistaken for athletic ability and skill development. Instead, the IMTF-S can be used as part of a comprehensive training program for athletes. It can be used as a baseline measure of mental toughness, from which a customized mental skills training program can be developed to enhance the four identified factors of mental toughness (Motivation, Identification, Negation, and Determination) over the course of a specified training period. The instrument can then be administered again to reveal changes and improvement in the mental toughness score. Due to the nature of sport, as athletes become more successful, they will likely experience greater levels of pressure, higher stakes, and more exposure. As a result, the IMTF-S can be utilized at multiple points throughout an athlete’s development in order to ensure that his/her mental toughness training is suitable for the demands and challenges that may be present or looming on the horizon.

5. Conclusion

The present research provides an overview of the development of a theoretically-sound instrument to measure mental toughness in sports. The psychometric properties of the instrument were evaluated by way of a powerful item response theory model, and results indicate the IMTF-S is a psychometrically-sound instrument capable of producing valid and reproducible measures of mental toughness in sports. The results of the current study provide further support and validation of the initial study conducted on the psychometric properties of the IMTF-S (Stonkus, 2011). It provides measureable and valuable insight on how an athlete experiences stress in sport. The IMTF-S has a number of important implications for both athletes and coaches, and additional use of the instrument is encouraged.

References

- Andrich, D. (1978). A rating formulation for ordered response categories. *Psychometrika*, 43, 561-573. <http://dx.doi.org/10.1007/bf02293814>
- Bandura, A. (Ed.). (1995). *Self-efficacy in changing societies*. New York: Cambridge University Press. <http://dx.doi.org/10.1017/cbo9780511527692>
- Bandura, A. (1977). Self-efficacy: Toward a unifying theory of behavioral change. *Psychological Review*, 84, 191-215. <http://dx.doi.org/10.1037/0033-295x.84.2.191>
- Bond, T. G., & Fox, C. M. (2007). Applying the Rasch Model. In *Fundamental measurement in the human sciences* (2nd ed.). Lawrence Erlbaum Associate. <http://dx.doi.org/10.4324/9781410614575>
- Crust, L. (2007). Mental toughness in sport: A review. *International Journal of Sport and Exercise Psychology*, 5, 270-290. <http://dx.doi.org/10.1080/1612197x.2007.9671836>
- Clough, P. J., Earle, K., & Sewell, D. (2002). Mental toughness: The concept and its measurement. In I. Cockerill (Ed.), *Solutions in sport psychology* (pp. 32-43). London: Thompson Publishing.

- Dweck, C. S. (2009). Mindsets: Developing talent through a growth mindset. *Olympic Coach*, 21(1), 4-7.
- Garnezy, N. (1993). Children in poverty: Resilience despite risk. *Psychiatry*, 56, 127-136.
- Goldberg, A. S. (1998). *Sports slump busting: 10 steps to mental toughness and peak performance*. Champaign, IL: Human Kinetics.
- Goldberg, A. S. (2005). *Sports slump busting: 10 steps to mental toughness and peak performance*. Coral Springs, FL: Llumina Press.
- Gould, D., Dieffenbach, K., & Moffet, A. (2002). Psychological talent and its development in Olympic champions. *Journal of Applied Sport Psychology*, 14, 177-210. <http://dx.doi.org/10.1080/10413200290103482>
- Jones, G., Hanton, S., & Connaughton, D. (2002). What is this thing called mental toughness? An investigation of elite sport performers. *Journal of Applied Sport Psychology*, 14, 205-218. <http://dx.doi.org/10.1080/10413200290103509>
- Kobasa, S. C. (1979). Stressful life events, personality, and health: An inquiry into hardiness. *Journal of Personality and Social Psychology*, 37, 1-11. <http://dx.doi.org/10.1037/0022-3514.37.1.1>
- Lazarus, R. S., & Folkman, S. (1984). *Stress, appraisal and coping*. New York: Springer. <http://dx.doi.org/10.1017/s0141347300015019>
- Linacre, J. M. (2015). *Correlations: Point-biserial, point-measure, residual*. Available at <http://www.winsteps.com/winman/correlations.htm>.
- Linacre, J. M. (2015). *WINSTEPS®* (Version 3.81). Computer Software. Beaverton, OR: Winsteps.com.
- Linacre, J. M. (2002). Optimizing rating scale category effectiveness. *Journal of Applied Measurement*, 3(1), 85-106.
- Loehr, J. E. (1986). *Mental toughness training for sport. Achieving athletic excellence*. Lexington, MA: Stephen Greene.
- Loehr, J. E. (1995). *The new mental toughness training for sports*. New York: Plume.
- Loehr, J. E. (1982). *Athletic excellence: Mental toughness training for sports*. New York: Plume.
- Mack, M. G., & Ragan, B. G. (2008). Development of the Mental, Emotional, and Bodily Toughness Inventory in collegiate athletes and non-athletes. *Journal of Athletic Training*, 43(2), 125-132. <http://dx.doi.org/10.4085/1062-6050-43.2.125>
- Messick, S. (1989). Validity. In R. L. Linn (Ed.), *Educational Measurement* (3rd ed., pp. 13-103). New York: Macmillan.
- Middleton, S. C., Marsh, H. W., Martin, A. J., Richards, G. E., & Perry, C. (2004a). *Discovering mental toughness: A qualitative study of mental toughness in elite athletes*. http://self.uws.edu.au/Conferences/2004_Middleton_Marsh_Martin_Richards_Perry.pdf
- Middleton, S. C., Marsh, H. W., Martin, A. J., Richards, G. E., & Perry, C. (2004b). *Developing the mental toughness inventory (MTI)*. Self Research Centre Biannual Conference, Berlin. http://self.uws.edu.au/Conferences/2004_Middleton_Marsh_Martin_Richards_Perry.pdf
- Middleton, S. C., Marsh, H. W., Martin, A. J., Richards, G. E., Savis, J., Perry, C., & Brown, R. (2004). The psychological performance inventory: Is the mental toughness test tough enough? *International Journal of Sport Psychology*, 35, 91-108.
- Norris, E. K. (1999). *Epistemologies of champions: A discursive analysis of champions' retrospective attributions: Looking back and looking within*. Michigan: Michigan University Microfilms International.
- Seligman, M. E. P. (2006). *Learned Optimism: How to change your mind and your life*. New York: Vintage.
- Sheard, M., Golby, J., & van Wersch, A. (2009). Progress toward construct validation of the Sports Mental Toughness Questionnaire (SMTQ). *European Journal of Psychological Assessment*, 25(3), 186-193. <http://dx.doi.org/10.1027/1015-5759.25.3.186>
- Stonkus, M. A. (2011). *The development and validation of the Inventory of Mental Toughness Factors in Sport (IMTF-S)* (Doctoral dissertation). Retrieved from ProQuest Dissertations & Theses (Publication number 3463253).

- Smith, R. M. (2000). Fit analysis in latent trait measurement models. *Journal of Applied Measurement, 1*(2), 199-218.
- Thelwell, R., Weston, N., & Greenless, I. (2005). Defining and understanding mental toughness within soccer. *Journal of Applied Sport Psychology, 17*, 326-332. <http://dx.doi.org/10.1080/10413200500313636>
- Williams, R. M. (1988). The U.S. Open character test: Good strokes help, but the most individualistic of sports is ultimately a mental game. *Psychology Today, 22*, 60-62.
- Wright, B. D., & Linacre, J. M. (1994). Reasonable mean-square fit values. *Rasch Measurement Transactions, 8*, 370.
- Wright, B. D., & Masters, G. N. (1982). *Rating scale analysis: Rasch Measurement*. Chicago: MESA Press.
- Wright, B. D., & Stone, M. (1999). *Measurement Essentials* (2nd ed.). Wilmington, DE: Wide Range, Inc.

Copyrights

Copyright for this article is retained by the author(s), with first publication rights granted to the journal.

This is an open-access article distributed under the terms and conditions of the Creative Commons Attribution license (<http://creativecommons.org/licenses/by/3.0/>).