Current Psychology
https://doi.org/10.1007/s12144-024-06562-9

Evaluation of the 10-item Mental Toughness Questionnaire (MTQ10): cross-cultural assessment and scrutiny of method effects

Andrew Denovan[1] · **Neil Dagnall**[2] · **Elena Artamonova**[2] · **Kostas A. Papageorgiou**[3,4]

Accepted: 14 August 2024

Abstract
The 10-item Mental Toughness Questionnaire (MTQ10) is an easy to administer, global measure of mental toughness. Previous analysis established that the MTQ10 was psychometrically superior to the preceding, widely used, 18-item version. Nevertheless, the MTQ10 is potentially undermined by a method effect arising from the inclusion of negatively phrased items. Noting this, the present paper further assessed the measurement properties of the MTQ10 via cross-cultural comparisons. This involved assessing the factor composition in samples from the UK ($N = 596$), Greece ($N = 1230$), and Italy ($N = 425$). Rather than a strict method effect, analyses found effects specific to pairs of negative items. The emergent model demonstrated partial invariance across countries. National variations in mental toughness scores were ascribable to societal differences. Convergent validity was demonstrated using theoretically related variables (Dark Triad and Life Satisfaction). Overall, results supported the use of the MTQ10 as a global mental toughness measure. Additionally, outcomes suggested that further cross-cultural comparison would provide useful insights into the nature of mental toughness.

Keywords Correlated uniqueness · Cross-cultural research · Measurement invariance · Mental toughness · MTQ10

Introduction

Contemporary delimitations of mental toughness were informed by seminal work with elite sportspersons (Loehr, 1982, 1986, 1995). This concluded that success in competition situations was associated with possession of a psychological edge/positive mindset (Kuan & Roy, 2007). Particularly, it indicated that mental toughness facilitated maintenance of high-performance levels, even when athletes were faced with challenges, barriers, and adversity (Connaughton et al., 2008b). Loehr (1986) attributed the enabling powers of mental toughness to the fact that the construct was allied to controlled, disciplined thought. Accordingly, during times of pressure, mentally tough individuals can remain relaxed, calm, and energized by increasing their flow of positive energy (Kuan & Roy, 2007).

Noting these attributes, theorists advanced the notion that mental toughness was a psychological construct that reflected the capacity to endure and manage stress effectively in order to perform well/maximally (Drinkwater et al., 2019). Despite being grounded within and highly pertinent to sporting environments (e.g., Dagnall et al., 2021; Mojtahedi et al., 2023; Wheatley et al., 2023), researchers consequently applied mental toughness to a range of everyday settings (i.e., occupational, Marchant et al., 2009; education, Denovan et al., 2021; and health, Levy et al., 2006). This occurred because investigators noted that these real-world domains, while externally different, inherently placed similar psychological pressures and demands on individuals. In addition to applicability, academic interest in mental toughness increased due to the construct's perceived explanatory power. Explicitly, theorists could conceptually accommodate mental toughness within models alongside related variables such as stress, coping, motivation, and self-esteem (Denovan et al., 2023). Moreover, mental toughness

✉ Andrew Denovan
a.m.denovan@ljmu.ac.uk

1 School of Psychology, Liverpool John Moores University, Liverpool, UK

2 Department of Psychology, Manchester Metropolitan University, Manchester, UK

3 School of Psychology, Queen's University Belfast, Belfast, UK

4 Department of Psychology, Neapolis University Pafos, Paphos, Cyprus

Published online: 20 September 2024

intuitively linked to health behaviors and outcomes (Bahmani et al., 2016).

The 10-item Mental Toughness Questionnaire (MTQ10)

Acknowledging the theoretical importance of mental toughness and the practical constraints of assessing the construct alongside multiple factors, researchers developed the brief, 10-item Mental Toughness Questionnaire (MTQ10) (Dagnall et al., 2019; Papageorgiou et al., 2018). The MTQ10 is an abridged form of the 48-item Mental Toughness Questionnaire (MTQ48) (Clough et al., 2002). The MTQ48 originated from Clough's mental toughness model (Clough et al., 2002). Clough drew heavily on the construct of hardiness (Kobasa, 1979) and operationalized mental toughness as a resource that guards against and counters stress (Crust & Keegan, 2010). Hence, Clough et al. (2002) created their 4 C model by adding Confidence to the hardiness sub-concepts of Control, Commitment, and Challenge. The inclusion of Confidence was designed to accommodate sport-related responses (Crust & Clough, 2005; Gerber et al., 2013).

Within the 4 C model, dimensions assess independent but correlated subcomponents: Control (Life and Emotion) indexes the inclination to feel influential and manage anxieties; Commitment refers to deep involvement in pursuing objectives, regardless of problems; Challenge denotes the propensity to perceive threats as self-development opportunities; and Confidence (in Abilities and Interpersonal) represents belief in self-worth, and the capacity to effectively navigate social interactions and settings. The subdivision of Control and Confidence led some investigators to describe the model as the 6Cs.

Although mental toughness research lacks conceptual clarity due to the existence of myriad definitions, the 4/6 C model is commensurate with expositions that denote mental toughness as a set of resources that enable task execution (see Gucciardi, 2017) and promote positive mental health (Drinkwater et al., 2019; Mojtahedi et al., 2020). Thus, the 4/6 C model aligns broadly with Gucciardi's (2017) contemporary definition of mental toughness as a flexible, effective, and purposeful psychological resource that facilitates and safeguards the pursuit of goals. This classification is important because it arose from a comprehensive overview of the conceptual progress of mental toughness.

The main problems with the 4/6 C model are structural stability and assumptions about the dispositional nature of mental toughness. Though the model is well recognized, widely applied, and endorsed by research (e.g., Perry et al., 2015, 2021) some attempts to reproduce the factorial structure have resulted in discrepancies (e.g., unsatisfactory model fit and context variations among subsamples; Gucciardi et al., 2013). Noting this, Gucciardi (2017) posits that mental toughness is a one-dimensional (as opposed to multifactorial), state-like construct. These suppositions, however, are contested (e.g., Perry et al., 2021).

Regarding contextual variations, Strycharczyk and Clough (2014) conceptualise mental toughness as a 'plastic' (trainable/malleable) trait. This delineation aligns with the view that mental toughness possesses a genetic basis (Horsburgh et al., 2009), and that improvements are most effectively achieved by concentrating on Commitment and Control, which demonstrate lowest heritability (Horsburgh et al., 2009). Notwithstanding these issues, the definition of Clough et al. (2002) concurs with the prevailing opinion that mental toughness is a complex concept, which aids the ability to cope with pressures/challenges. Thus, the MTQ48 is a valuable instrument since it indexes a breadth of construct relevant content.

The development of the MTQ10 was necessary for several reasons. The initial motivation derived from the observation that the original shortened version of the MTQ48, the 18-item Mental Toughness Questionnaire (MTQ18), featured in research despite limited psychometric substantiation. Indeed, the author's principal rationale for using the MTQ18 was grounded in the psychometric qualities of the parent measure. The need for a concise instrument arises from the fact that investigators require a measure that can be expediently administered within large test batteries. Shorter scales are also desirable when limited testing time exists, or accessibility is a concern. Explicitly, when the respondent group has cognitive limitations (e.g., youthful respondents have briefer attention spans and are susceptible to distraction).

Noting the lack of direct psychometric evidence for the MTQ18 and the need for an easily administrable and freely available short measure, the MTQ10 was created (Papageorgiou et al., 2018). To ensure that each of the 4Cs were sampled, the item pool comprised 12 items (three from each factor). Confirmatory factorial analysis indicated that two items (one from Challenge and Commitment) loaded weakly on a general component. These were subsequently removed. Further cross-lagged assessment, using two waves of data, indicated the MTQ10 possessed temporal stability (Papageorgiou et al., 2018).

Dagnall et al. (2019) further examined the factorial structure of the MTQ10 and MTQ18 scales by assessing fit with single (1-factor and 1-factor correlated) and four factor (correlated and bifactor) solutions. Analysis found that the single factor models provided the best fit. Comparison with the MTQ18 (Clough et al., 2002) revealed that the MTQ10 was a superior global index, possessing greater factor loadings alongside stronger data fit. Additionally, concurrent validity

tests indicated that the MTQ10 was a more robust prognosticator of life satisfaction (well-being). Gender invariance (configural, metric, and scalar) existed for both measures.

Accordingly, Dagnall et al. (2019) concluded that although the MTQ18 was a psychometrically adequate scale, the MTQ10 was a superior global mental toughness measure. Notwithstanding these outcomes, the best solution was less than optimal since the presence of negatively phrased items necessitated the correlating of error terms between two item pairs (DiStefano & Motl, 2006). Ensuing analysis of the MTQ10 by Kawabata et al. (2021) supported these findings. Although, Kawabata et al. (2021) subsequently failed to correlate specific error terms.

Test developers typically incorporate negatively phrased items within surveys because, by virtue of being more cognitively demanding, they interrupt the flow of answering. This is advantageous within scales as it encourages participants to reflect upon item content. The concomitant benefit to the test developer is a reduction in automatic responding, which can give rise to bias (Podsakoff et al., 2012). Despite possessing potential merit, some theorists question the usefulness of negatively worded items, contending that they do not actually reduce response bias and produce data contaminated by respondent inattention and confusion (Sonderen et al., 2013).

Acknowledging these concerns, the need to correlate negative items within the MTQ10 may reflect the presence of systematic measurement error arising from shared method (i.e., item polarity) and/or content (i.e., question substance/phrasing). For instance, negatively worded items may index common factors such as poor anxiety management (e.g., item 2, 'I tend to worry about things well before they actually happen'; and item 7, 'When I make mistakes, I usually let it worry me for days after') and lack of purpose/direction (i.e., item 3, 'I usually find it hard to summon enthusiasm for the tasks I have to do'; and item 6, ' "I just don't know where to begin" is a feeling I usually have when presented with several things to do at once').

Though further research is required to determine the precise reason for the negative wording effect within the MTQ10, it is suggestive of a method factor. A method effect exists when the direction of item wording influences responses. This is problematic because the underlying assumption is that items, regardless of directionality, assess the same construct (Marsh, 1996). Thus, method effects contaminate content under investigation (i.e., limit factorial validity). This is concerning because it can obstruct perception of the construct under investigation, and influence data interpretation. For example, method variance can inflate or suppress the association between two scales. Such outcomes lead to inaccurate interpretation of relationship significance (DiStefano & Motl, 2006). Thus, the presence of a method effect can skew understanding of a construct/measure and allied data.

The present study

A principal objective of the current paper was to determine whether a method effect was present within the MTQ10. Correspondingly, the authors assessed a series of structural solutions: a single-factor, correlated methods (CM), a correlated uniqueness (CU), and a CU model based on Dagnall et al. (2019) (see Fig. 1). These models emerged from a multitrait-multimethod (MTMM) framework, which focuses on separating construct-specific information (mental toughness) from method effects (Lindwall et al., 2012). The CU model included correlations among residuals of uniqueness (i.e., negative items). For the CU model based on Dagnall et al. (2019), correlations occurred among residuals of items 2 and 7, and 3 and 6. In contrast, the CM model incorporated a latent method effect factor.

A MTMM framework was appropriate because it enabled the researchers to perform a construct (mental toughness) vs. method used to acquire responses comparison (DiStefano & Motl, 2006). CM permits fragmentation of variance into trait, method, and error via modelling the method as a latent variable, whereas CU correlates disturbances among negative items. Support for a method factor specifies the presence of a construct, separate from mental toughness, that requires additional exposition. This could be trait-based and/or reflective of a pervasive approach to responding to the MTQ10 (e.g., a response style) (Kam & Sun, 2020).

Moreover, evidence of a method factor necessitates additional analysis to identify the origin of the factor and specify its relationship with other variables (Lima & Souza, 2019). Conversely, support for a CU model advises the need to control for specific variance (arising from the negative items) during analyses to achieve a clearer representation of the construct (DiStefano & Motl, 2006). Hence, the use of MTMM models provides insight into the nature of the 'cognitive speed bump' present within the MTQ10. Accordingly, it was necessary for analysis to establish whether models comprising method components or correlated errors fitted data better than a single-factor model without constraints.

The other principal objective of this investigation was to evaluate cross-cultural applicability of the MTQ10. To achieve this, comparisons were made between independent samples from nations with different cultural orientations (i.e., United Kingdom, Greece, and Italy) to determine if the factor and latent mean composition of the MTQ10 was invariant across nationalities. A concomitant practical outcome was the production of different language versions of the MTQ10.

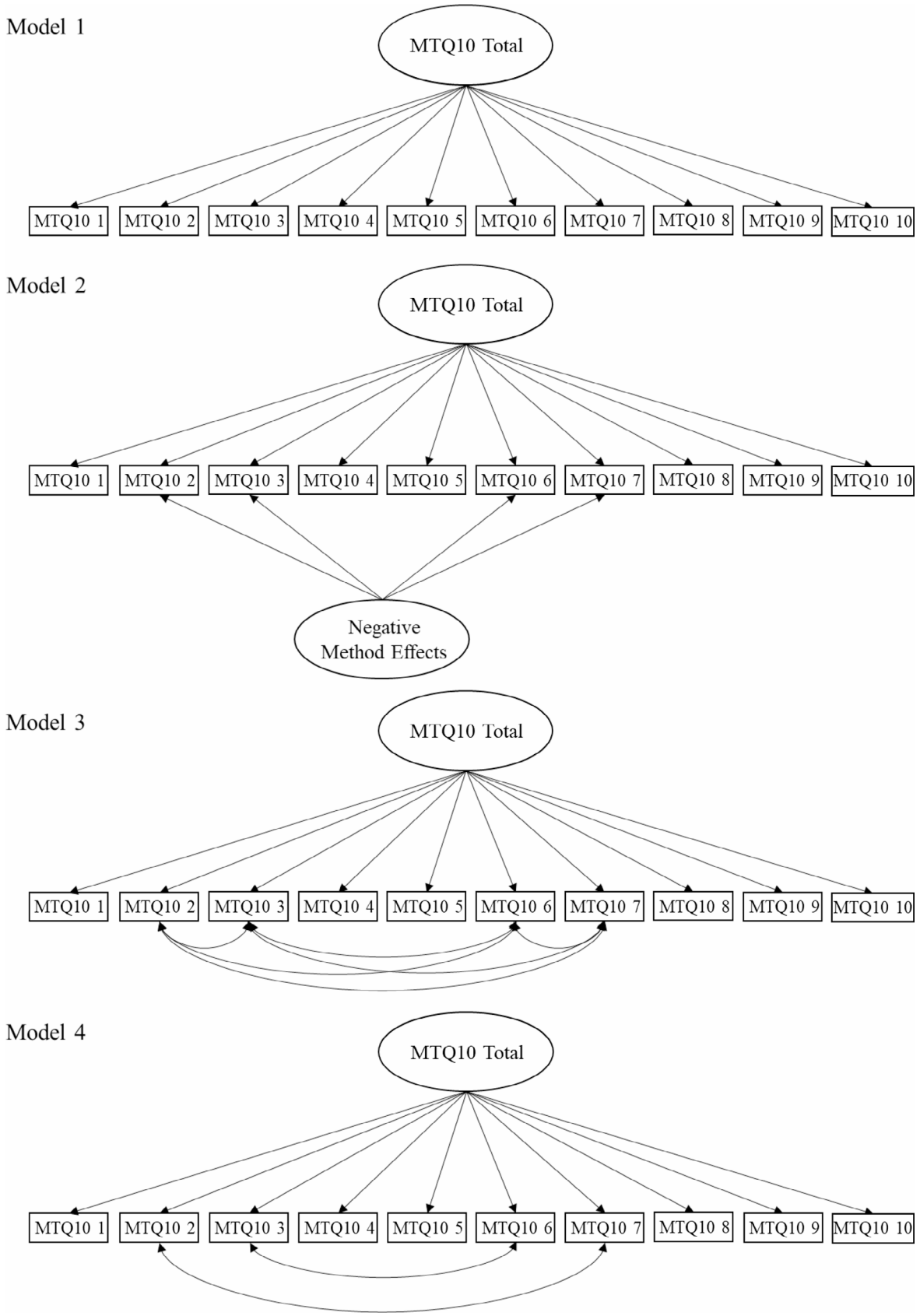

Fig. 1 Competing MTQ10 models

Scrutiny of cross-cultural MTQ10 applicability is important because it determines measurement model equivalency across populations (Cowden, 2018). Developing language-specific versions of the MTQ10 enables cross-national comparisons (Behr & Sha, 2018), which due to the newness of the scale are currently lacking. The UK sample was selected as the representative group because the original MTQ48 was developed and validated in a UK context. Italy and Greece were targeted because mental toughness is a construct of interest within these countries (Papageorgiou et al., 2018; Stamatis et al., 2021). Moreover, Greece is a collectivist society (Hofstede, 2001). Accordingly, there may be unique variation in the precedence given to individual traits such as mental toughness, due to more of an emphasis on group objectives and roles (Triandis et al., 1995). Italy, however, reflects more of an individualistic society (Burton et al., 2021). Thus, direct comparisons can be made with the UK (also individualistic).

External (convergent) validity was assessed by comparing the MTQ10 with relatively context-free constructs that have been examined alongside mental toughness. Explicitly, the Dark Triad (comprising Machiavellianism, subclinical narcissism, and subclinical psychopathy) and life satisfaction. Research indicates that the MTQ10 (if accurately approximating mental toughness) should correlate positively and significantly with narcissism and life satisfaction (e.g., Dagnall et al., 2019; Papageorgiou et al., 2019). Meanwhile, associations with Machiavellianism and psychopathy should be negative and weak (< 0.10) (Papageorgiou et al., 2017).

Assertions regarding mental toughness and the Dark Triad arose from an established literature base (e.g., Szabo et al., 2022). This reports that, due to shared attributes, mental toughness is expected to correlate positively with narcissism. Explicitly, individuals high in subclinical narcissism possess a strong belief in their abilities (Onley et al., 2013). This confidence is a shared quality between mentally tough and narcissistic individuals (i.e., the Confidence facet of mental toughness correlates the most strongly with subclinical narcissism).

In addition, narcissists typically possess a strong commitment to their objectives, while disregarding the consequences, which also characterises mentally tough individuals (Sabouri et al., 2016). Negative associations regarding mental toughness and Machiavellianism are suggested to occur due to individuals higher in Machiavellianism possessing a cynical worldview (e.g., in relation to the link between hard work and achievement), which is incompatible with the drive towards commitment and achievement among mentally tough individuals (Jones & Paulhus, 2014). A negative relationship should exist between psychopathy and mental toughness because psychopathic individuals lack the self-control that typifies mentally tough individuals (Onley et al., 2013).

Regarding life satisfaction, research has frequently reported positive correlations with mental toughness. This has been determined using the MTQ48, MTQ18, and MTQ10 (Clough et al., 2002; Dagnall et al., 2019; Gerber et al., 2013). Life satisfaction is representative of positive psychology characteristics, which consistently associate with higher mental toughness (e.g., self-esteem, Earle, 2006; optimism, Nicholls et al., 2008). Moreover, mental toughness reflects a resilience-based construct, which enables individuals to adjust more effectively to situations. Positive adjustment and functioning are embedded within life satisfaction (Gerber et al., 2013).

Materials and methods

Respondents

This study used three independent national samples recruited via adverts on social networks and word of mouth. Minimum sample size requirements were determined using a simulation in accordance with the N: q criteria (Schumacker & Lomax, 2015). This indicated that > 360 respondents were required for each sample. All data collection occurred online. Initial data screening removed cases with missing data and/or z-scores > 3.29 or < -3.29 *SD*s from the mean (19 UK, 13 Greece, and 6 Italy) (Tabachnick & Fidell, 2013). Final samples comprised: UK 596 participants, 417 females and 179 males (*M*age = 27.97, *SD* = 11.23, range of 18 to 71); Greece 1230 participants, with 854 females and 376 males (*M*age = 35.17, *SD* = 13.02, range of 18 to 86); and Italy 425 participants, 311 females and 114 males (*M*age = 27.55, *SD* = 10.74, range of 18 to 69).

Of the UK sample, 18%, 27%, and 55% occupied high school, college, and higher education qualifications. For the Greece sample, 5%, 34%, and 61% fell into these respective categories, as did 2%, 49%, and 49% of the Italy sample.

Instrument translation

To facilitate data collection in Greece and Italy, two native speaking external collaborators (i.e., not co-authors) performed forward translation of the MTQ10, the Short Dark Triad (for Greece only), and the Satisfaction with Life Scale (for Italy only) given language specific versions of these measures were not available. Then, a co-author and an external collaborator proficient in English, whose native languages are Greek and Italian respectively, assessed inadequate expressions/concepts and performed back-translation. The original and the back translated versions exhibited

no major differences. Back translation is a popular method within cross-cultural research, which is effective for survey adaptation (Son, 2018).

Measures

The 10-item Mental Toughness Questionnaire (MTQ10)

Within the MTQ10 items appear as statements (e.g., 'I generally feel that I am a worthwhile person'). Participants use a five-point Likert scale (1 = Strongly Disagree to 5 = Strongly Agree) for responding. Computation of item totals generates a score between 10 and 50. Lower scores indicate mental sensitivity and higher scores mental toughness. The MTQ10 has evidenced satisfactory reliability using a range of methods (composite, alpha, and test-retest) (Dagnall et al., 2019; Papageorgiou et al., 2018). In this study, the MTQ10 demonstrated good alpha and omega reliability for the UK sample ($\alpha = 0.83$, $\omega = 0.93$), and good to satisfactory for the Greek ($\alpha = 0.79$, $\omega = 0.88$), and Italian ($\alpha = 0.78$, $\omega = 0.91$) samples.

The short Dark Triad (SD3)

The SD3 (Jones & Paulhus, 2014) captures Machiavellianism, narcissism, and psychopathy using 27-items (e.g., 'Many group activities tend to be dull without me'). Participants utilize a five-point Likert scale (1 = completely disagree to 5 = completely agree) for responding to each statement. Somma et al. (2019) translated the scale to Italian and reported acceptable reliability. In this study, $\alpha = 0.70$, $\omega = 0.70$, $\alpha = 0.73$, $\omega = 0.73$, and $\alpha = 0.73$, $\omega = 0.70$ for Machiavellianism, narcissism, and psychopathy for the UK sample. For the Greece sample, $\alpha = 0.75$, $\omega = 0.75$ (Machiavellianism), $\alpha = 0.69$, $\omega = 0.70$ (narcissism), and $\alpha = 0.76$, $\omega = 0.77$ (psychopathy). For Italy, $\alpha = 0.79$, $\omega = 0.80$ (Machiavellianism), $\alpha = 0.71$, $\omega = 0.72$ (narcissism), and $\alpha = 0.75$, $\omega = 0.76$ (psychopathy).

The satisfaction with Life Scale (SWLS)

The SWLS (Diener et al., 1985) assesses contentment with life via statements (e.g., "In most ways, my life is close to my ideal") focusing on global cognitive appraisal. The scale items are accompanied by a seven-point Likert format (ranging from 1 = strongly disagree to 7 = strongly agree). The SWLS indexes satisfactory psychometric properties (construct validity, internal and test–retest reliability; Pavot & Diener, 2008) and has been successfully translated to Greek (Galanakis et al., 2017). Internal reliability of the SWLS was satisfactory in the present investigation for the UK ($\alpha = 0.89$, $\omega = 0.88$), Greece ($\alpha = 0.87$, $\omega = 0.87$), and Italy ($\alpha = 0.85$, $\omega = 0.85$) samples.

Procedure and ethics

Following advertisements on social networks, interested participants received information about the study. All participants consented to participate and were contacted with a password alongside a link to complete the online survey. Forced response was used for the survey (i.e., progression to a subsequent page only occurred once participants had answered all items on a proceeding page). This was necessary to limit missing information in participant responses. After completing the study, all respondents were debriefed.

The project received ethical authorisation from the appropriate institutional review board. The identity of the review board is withheld to preserve anonymity. The study procedures conformed to the ethical guidelines of the institution and concurred with the 1964 Helsinki declaration.

Analysis

To evaluate the psychometric performance of the MTQ10 across countries (i.e., UK, Greece, and Italy), analysis using Mplus v7 assessed the four CFA models (see Fig. 1). These were selected to replicate preceding MTQ10 investigations (i.e., Dagnall et al., 2019), and examined potential method effects arising from reverse-worded items. Model 1 represented a one-factor solution. Model 2 was a two-factor CM model, with a mental toughness factor and an additional factor signifying wording effects among the four negative items. Model 3 paralleled Model 2 but applied a CU model. Model 4 also applied a CU model. Within this, wording effects were depicted by specifying correlations among error terms for the negatively phrased items (see Dagnall et al., 2019). This was the superior MTQ10 solution advanced by Dagnall et al. (2019).

Judgement of model fit involved consultation of fit indices. Specifically, chi-square, Comparative Fit Index (CFI), Root-Mean-Square Error of Approximation (RMSEA) and Standardized Root-Mean-Square Residual (SRMR). CFI ≥ 0.90, SRMR ≤ 0.08 and RMSEA ≤ 0.08 reflected satisfactory fit (Bentler, 2007; Chen et al., 2008). Marginal fit represented CFI ≥ 0.89, SRMR ≤ 0.10 and RMSEA ≤ 0.10 in accordance with several published papers (e.g., Dagnall et al., 2016; Rogers et al., 2005; Sola et al., 2018).

Successively, invariance tests (configural, metric and scalar) ensued comparing the three countries. Findings from simulation studies (Rutkowski & Svetina, 2014) recommended the subsequent cutoffs: A CFI change ≤ 0.02 alongside an RMSEA variation of ≤ 0.01 to conclude loading invariance; a change in CFI and RMSEA ≤ 0.01 to conclude

equivalence of intercepts. Latent factor means among the invariance groups were subsequently compared prior to assessing convergent validity.

Results

Data screening

For all national samples, univariate skewness and kurtosis tests reported satisfactory values between −2.0 and +2.0 (Field & Miles, 2010). However, multivariate kurtosis (Mardia's *b2p*) and skewness (Srivastava's *b1p*) indicated departure from multivariate normal distribution. Specifically, for kurtosis: UK sample = 13.45, $p < .001$; Greece sample = 18.01, $p < .001$; Italy sample = 6.44, $p < .001$. For skewness: UK sample = 22.01, $p = .015$; Greece sample = 50.78, $p = .015$; Italy sample = 24.86, $p = .006$. Therefore, an estimation method that produces robust standard errors and estimates under non-normality (i.e., MLR estimation) was necessary (Marsh et al., 2013).

Confirmatory factor analysis

Model 1 reported unsatisfactory CFI, marginal RMSEA, and good SRMR across countries (Table 1). All items loaded significantly for all samples, but modification indices recommended correlating errors among the negative items of 2 and 7, and 3 and 6. Model 2 revealed good fit for the UK and Greece samples. For the Italy sample, good RMSEA and SRMR, but marginal CFI existed. Scrutiny of factor loadings suggested a non-significant loading of item 3 on the method factor for the UK sample ($p = .076$). Similarly, no items loaded significantly onto the method factor in the Italian sample. All items loaded significantly for the Greece sample. However, for the Greece sample a greater average loading existed for the general factor negative items in comparison with the method factor items (0.38 vs. 0.36). Table 2 includes factor loadings for competing models alongside loadings for Model 2 method factor.

Table 1 Fit indices for competing country specific MTQ10 solutions

Model	χ^2	df	CFI	SRMR	RMSEA (90% CI)
Model 1					
UK	193.18**	35	0.88	0.05	0.08 (0.07-0.09)
Greece	360.65**	35	0.86	0.05	0.08 (0.07-0.09)
Italy	138.15**	35	0.86	0.05	0.08 (0.06-0.09)
Model 2					
UK	116.98**	31	0.93	0.03	0.06 (0.05-0.08)
Greece	244.59**	31	0.91	0.04	0.08 (0.06-0.08)
Italy	109.80**	31	0.89	0.04	0.07 (0.06-0.09)
Model 3					
UK	75.33**	29	0.96	0.03	0.05 (0.03-0.06)
Greece	202.18**	29	0.93	0.03	0.07 (0.06-0.07)
Italy	90.87**	29	0.91	0.04	0.07 (0.05-0.08)
Model 4					
UK	106.48**	33	0.94	0.04	0.06 (0.04-0.07)
Greece	252.26**	33	0.91	0.04	0.07 (0.06-0.08)
Italy	101.16**	33	0.91	0.04	0.07 (0.05-0.08)

**χ^2 significant at $p < .001$

The CU model (Model 3) demonstrated good fit in each sample. However, non-significant correlations existed among error terms between items 2 and 3 (UK $p = .076$), 2 and 6 (Italy $p = .146$), 3 and 7 (UK $p = .120$; Greece $p = .172$; Italy $p = .655$), and 7 and 6 (Italy $p = .067$). For Model 4 (the CU model based on Dagnall et al., 2019), CFA indicated satisfactory fit for the UK, Greek, and Italian samples. All items loaded significantly for all samples, and correlations among error terms concerning negative items 2 and 7, and 3 and 6 were significant (all $p < .001$). Comparatively, this model demonstrated better data-model fit than Model 1 and 2 for all samples apart from the Greece sample. However, for this sample Model 4 was judged as superior on the basis Model 2's RMSEA was greater, and the method factor negative items exhibited a lower average loading than the general factor.

Moreover, Model 3 demonstrated the best data-fit for each sample. Conversely, this model evidenced non-significant correlated uniqueness for all negative item pairs but 2 and 7, and 3 and 6. In this instance, it is possible that the presence of additional correlated uniqueness improved fit, but not in a meaningful way due to the non-significant correlations. This is supported by the MI indices from Model 1 recommending correlating uniqueness among negative items 2 and 7, and 3 and 6 only. Accordingly, Model 4 represented the superior model based on model fit *and* the significance of parameters. Model 4 factor loadings (Table 2) for the UK sample ranged from 0.47 to 0.74, with an average of 0.57. For the Greek and Italian samples, factor loadings ranged from 0.36 to 0.75 and 0.36 to 0.73 respectively, comprising average loadings of 0.53 and 0.51.

Invariance analysis

Tests of invariance included applying Model 4 (the best fitting solution) to the data across nations. The configural model (Table 3) revealed good fit (CFI = 0.91, RMSEA = 0.07). This infers that the quantity of MTQ10 latent factors is stable across groups. Comparison of configural (form) and metric (factor structure) models evidenced an acceptable CFI difference of 0.01 and no RMSEA difference. Assessing metric vs. scalar models revealed an unacceptable change in CFI (0.06). Relaxing the intercepts for items 3, 4, and 5 resulted in a CFI difference of 0.01 and no difference in

Table 2 Standardized factor loadings across model and nation

MTQ10 item	Model 1			Model 2			Model 3			Model 4		
	UK	GR	IT	UK (MFL)	GR (MFL)	IT (MFL)	UK	GR	IT	UK	GR	IT
1 Even when under considerable pressure I usually remain calm	0.49	0.66	0.54	0.49	0.66	0.53	0.50	0.66	0.54	0.50	0.67	0.54
2 I tend to worry about things well before they actually happen	0.52	0.39	0.41	0.44 (0.51)	0.34 (0.41)	0.36 (0.68)	0.45	0.34	0.37	0.47	0.36	0.38
3 I usually find it hard to summon enthusiasm for the tasks I have to do	0.56	0.43	0.37	0.52 (0.21)	0.41 (0.16)	0.36 (0.11)	0.52	0.40	0.35	0.53	0.41	0.36
4 I generally cope well with any problems that occur	0.65	0.73	0.70	0.67	0.75	0.72	0.67	0.75	0.73	0.67	0.75	0.73
5 I generally feel that I am a worthwhile person	0.64	0.39	0.51	0.66	0.41	0.53	0.67	0.41	0.53	0.66	0.41	0.53
6 "I just don't know where to begin" is a feeling I usually have when presented with several things to do at once	0.56	0.43	0.45	0.49 (0.37)	0.39 (0.32)	0.43 (0.11)	0.47	0.38	0.42	0.51	0.41	0.43
7 When I make mistakes, I usually let it worry me for days after	0.58	0.44	0.39	0.50 (0.51)	0.39 (0.53)	0.34 (0.38)	0.51	0.40	0.35	0.54	0.42	0.36
8 I generally feel in control	0.72	0.59	0.62	0.74	0.60	0.62	0.74	0.60	0.62	0.74	0.60	0.62
9 I am generally able to react quickly when something unexpected happens	0.45	0.61	0.51	0.47	0.62	0.52	0.47	0.63	0.52	0.47	0.62	0.52
10 I generally look on the bright side of life	0.60	0.61	0.60	0.61	0.61	0.59	0.61	0.61	0.59	0.61	0.61	0.60

GR Greece, *IT* Italy, *MFL* Method factor loading

RMSEA. This amendment resulted in partial scalar invariance of the MTQ10 across countries.

Latent mean comparison

Since partial scalar invariance existed across nations, an assessment of latent means occurred (Heim et al., 2017). The UK was the reference group, and its latent means were constrained to zero (Hong et al., 2003). Compared to the UK, significantly higher MTQ10 means existed for Greece ($M = 0.07, p = .045$). Moreover, Italy evidenced significantly lower MTQ10 latent means ($M = -0.23, p < .001$). A small effect size (Cohen's d) occurred in each instance (Greece vs. UK $d = 0.12$, Italy vs. UK $d = 0.38$).

Convergent validity

Comparison of the MTQ10 with SD3 and SWLS measures (Table 4) revealed positive associations with narcissism and life satisfaction. Moreover, the MTQ10 corelated weakly and negatively with Machiavellianism and psychopathy. These results were as expected. Correlation strength significantly differed for a minority of comparisons. Explicitly, Italy possessed greater correlations than Greece with regards to MTQ10 scores and Machiavellianism ($z = 2.62$, $p = .008$), psychopathy ($z = 2.24$, $p = .024$), and life satisfaction ($z = 3.63$, $p < .001$). The UK evidenced greater correlations than Greece concerning MTQ10 and narcissism ($z = 2.10$, $p = .034$), but weaker correlations than Italy for MTQ10 and life satisfaction ($z = 2.40$, $p = .016$).

Discussion

Using a MTMM framework, this study assessed the degree to which method effects specific to negatively keyed items were present with the MTQ10. This included evaluation of statistical models comprising a latent method factor and correlated errors among negatively worded items. Findings recommended correlating errors among negative items possessing similar content. The presence of a negative item method effect was not convincingly supported. Additional objectives included evaluating the cross-cultural applicability of the MTQ10, using UK, Greek, and Italian samples, and examining convergent validity in relation to theoretically aligned constructs (the Dark Triad and Life Satisfaction). Results supported partial invariance of the MTQ10 and revealed expected correlation patterns (i.e., positive associations with narcissism and life satisfaction, weak negative associations with Machiavellianism and psychopathy).

Analysis of potential method effects indicated greatest fit for Model 4, the correlated uniqueness (CU) model based

Table 3 Fit indices for MTQ10 invariance models across country (UK, Greece, and Italy)

Model	χ^2	df	CFI	CFI difference	SRMR	RMSEA (90% CI)	RMSEA difference
Configural	460.16**	99	0.92		0.04	0.07 (0.06-0.07)	
Metric	564.89**	117	0.91	0.01	0.07	0.07 (0.06-0.07)	No difference
Scalar	858.91**	135	0.84	0.06	0.10	0.08 (0.07-0.09)	0.02
Scalar (partial)	597.94**	131	0.90	0.01	0.08	0.07 (0.06-0.07)	No difference

χ^2 chi-square goodness-of-fit statistic, *df* degrees of freedom, *CFI* Comparative Fit Index, *SRMR* Standardized Root-Mean-Square Residual, *RMSEA* Root-Mean-Square Error of Approximation, *AIC* Akaike Information Criterion; **χ^2 significant at $p < .001$

Table 4 Correlations of the MTQ10 with SD3 subscales and the SWLS

Variable	1	2	3	4	5
UK sample					
1 MTQ10		−0.05	0.30**	−0.01	0.36**
2 SD3 Machiavellianism			0.33**	0.51**	−0.13**
3 SD3 narcissism				0.39**	0.17**
4 SD3 psychopathy					−0.19**
5 SWLS					
Greece sample					
1 MTQ10		−0.01	0.19**	−0.06*	0.32**
2 SD3 Machiavellianism			0.25**	0.39**	−0.08*
3 SD3 narcissism				0.32**	0.11**
4 SD3 psychopathy					−0.22**
5 SWLS					
Italy sample					
1 MTQ10		−0.16**	0.22**	−0.15**	0.49**
2 SD3 Machiavellianism			0.35**	0.56**	−0.09
3 SD3 narcissism				0.37**	0.16**
4 SD3 psychopathy					−0.20**
5 SWLS					

*$p < .05$; **$p < .001$

on Dagnall et al. (2019). This model, particularly due to the existence of non-significant correlations among several negative items in Model 3, was the most expedient solution. This reported good model fit, significant within-item error correlations, and good average factor loadings (from 0.51 to 0.57) across nations. The significant within-item error correlations occurred for items 2 and 7, and 3 and 6.

Furthermore, Model 1 results supported the need to correlate the error terms specified within Model 4. These findings suggest that a general negative item method effect was not present. However, an effect specific to pairs of the negative items occurred. This was attributable to similarities in wording/direction *and* content. Explicitly, items 2 ('I tend to worry about things well before they actually happen') and 7 ('When I make mistakes, I usually let it worry me for days after') assess worry-related phenomena. Items 3 ('I usually find it hard to summon enthusiasm for the tasks I have to do') and 6 ('"I just don't know where to begin" is a feeling I usually have when presented with several things to do at once') reference lack of volition/purpose. These components represent important inverse characteristics of mental toughness (sensitivity). Specifically, low resilience and control (Clough et al., 2002). Thus, although the presence of negative correlations reduced factor solution clarity, inclusion was conceptually necessary as these items index integral aspects of mental toughness. This approach contrasts with Kawabata et al. (2021), whose expedient solution was statistically rather than theoretically driven.

Regarding national samples, Model 4 fit was good for UK, Greek, and Italian data; all items loaded significantly. These findings indicate that structurally, the Greek and Italian translations are commensurate with the UK, English language version. Further analyses supported invariance at the factor (metric) level, and partial invariance (equivalence of intercepts, scalar). Cross-cultural non-invariance can occur due to differences in context and/or translation nonequivalence (International Test Commission, 2017). Specifically, translation differences. For instance, item 5 'I generally feel that I am a worthwhile person' translated in Italian as 'I generally feel like a worthy person', which possesses a slightly different meaning. Occurrences similar to this can impact factor loadings.

Compared to the UK, the Greek sample scored significantly higher on the MTQ10, and the Italian sample scored lower. Differences between Greece and the UK are interpretable using Hofstede's (2001) dimension of individualism vs. collectivism. Explicitly, Greece scores lower than the UK (35 vs. 89), reflecting a more collectivist culture. A lack of research has examined collectivism-individualism and mental toughness. However, evidence suggests that collectivist cultures promote aspects including social support (Goodwin & Plaza, 2000), which are important for mental toughness development (Connaughton et al., 2008a; Crust & Clough, 2011). The higher score for the UK vs. Italy is interpretable using Hofstede's (2001) uncertainty avoidance dimension. Italy scores higher than the UK (75 vs. 35), reflecting higher anxiety and stress levels, which may result in greater 'mental sensitivity'. Indeed, research consistently shows that stress is negatively associated with mental toughness (e.g., Gerber et al., 2018). These postulations require further research to distinguish the cultural factors that cultivate these differences.

Convergent validity indicated positive associations of the MTQ10 with narcissism and life satisfaction across nations. Furthermore, the MTQ10 corelated weakly and

negatively with Machiavellianism and psychopathy. These findings are consistent with preceding work (Dagnall et al., 2019; Papageorgiou et al., 2017, 2019). Particularly, narcissism has been consistently reported to index the strongest relationship of the Dark Triad traits with mental toughness (Onley et al., 2013). This is due to narcissism possessing similar qualities (e.g., high confidence). Furthermore, the SD3 (used in this study) typically captures grandiose features of narcissism, including high self-esteem and self-enhancement rather than antisocial aspects (Papageorgiou et al., 2023). This measurement inclination further suggests why positive associations frequently occur. Machiavellianism and psychopathy were (weakly) negatively associated with mental toughness, aligning with the findings of Szabo et al. (2022).

Szabo et al. (2022) suggested that these results may occur due to the SD3 capturing features that do not correspond strongly with mental toughness, including cynicism, antisocial behavior, and erratic lifestyle. Effect sizes/correlation intensities in this study were similar to Szabo et al. (2022) (who studied Dark Triad and mental toughness relationships with five independent samples) when using the criteria of Gignac and Szodorai (2016) for individual differences research (i.e., weak for mental toughness with Machiavellianism and psychopathy, strong for mental toughness and narcissism).

It is not obvious why a small quantity of significantly different associations existed across samples. However, these mainly occurred for measures that were translated in this study (Greece SD3, Italy SWLS). It is possible that stronger correlations existed between MTQ10 and convergent criteria for the Italy sample (vs. the UK and Greece) due to greater Uncertainty Avoidance. Particularly, higher anxiety and intolerance of uncertainty might amplify the significance of mental toughness and criteria including life satisfaction for this nation. This could be due to a greater awareness of stress as a feature of everyday life (Hofstede, 2001). Conversely, results existing mainly for the translated scales potentially indicate issues with capturing the constructs effectively. Perhaps additional work needs to be done when translating such measures in future (e.g., distribute to pilot samples).

Limitations and implications

The present investigation shares oft-cited limitations with other studies in this area. Notably, self-report measures are prone to concerns including social desirability and common-method variance (Podsakoff et al., 2003). This is particularly the case with mental toughness, which indexes performance advantages. Furthermore, samples comprised more women than men and evidenced differences in educational attainment.

Regarding partial invariance, difficulties in translation can arise at the content analysis and selection of equivalence stages. Indeed, psychological terms do not always have stable corresponding terms and rely upon the appropriate selection of equivalent expression. This can prove problematic, as replacements may possess pre-existing meanings and associations. Accordingly, when translating English psychological terms, Balygina and Ermolova (2018) suggest combining analysis of terminology with a consideration of contextual conditions. This is sometimes difficult since psychological meanings are already well established and not readily able to accommodate cultural variations and nuances.

An additional limitation concerns the reliance on classical test theory for examining the MTQ10's psychometric properties. A restriction of classical test theory is that differences in item difficulty and response scale functioning cannot be assessed. Disorders in response scale performance can exist, which negatively impact the psychometric integrity of a measure (Linacre, 2012). It would be useful for future research to assess the MTQ10 using techniques such as Rasch analysis, which enable scrutiny of item and response scale functioning. Indeed, related mental toughness measures (e.g., the Mental, Emotional, and Bodily Toughness Inventory) have used Rasch analysis to assist with scale construction and evaluation (Mack & Ragan, 2008).

Study limitations aside, this investigation had access to three large, representative national samples, and furthered understanding of the dependability and validity of the MTQ10. General support for convergent validity across countries adds further evidence of construct validity for the MTQ10, which is a crucial objective during scale development/evaluation (Grimm & Widaman, 2012). Importantly, future researchers can use the MTQ10 with greater confidence that the measure is capturing global mental toughness. Availability of a cross-language edition of the MTQ10 is anticipated to increase cross-cultural validity research on this measure alongside further work on the cultural forces that contribute to individual differences in mental toughness.

Conclusion

Overall, findings support the use of the MTQ10 as a brief, easy to administer general measure of mental toughness. Although significant linkages existed among negatively worded items, it is recommended to retain these items given they represent important inverse features of mental toughness and administer the measure in its entirety. In addition,

the MTQ10 performed well across three countries with the inclusion of specific correlated disturbances, with good support for its factor structure and loadings. The observation of instances of item non-invariance suggests that further work is required to produce equivalent foreign language statements. This is particularly important considering the original version of the MTQ10 (the MTQ48) was designed and validated using English-speaking participants.

Author contributions AD and ND designed the study and prepared the article. AD additionally prepared the data and completed the analysis. EA and KP edited, reviewed, and commented on drafts.

Funding This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Data availability Study data are available via figshare: https://figshare.com/s/1eb848e841bc3d33e227.

Declarations

Ethics approval Approval was obtained from the ethics committee of Queen's University Belfast. The procedures used in this study adhere to the tenets of the Declaration of Helsinki.

Consent Informed consent was obtained from all individual participants included in the research.

Competing interests The authors have no competing interests to declare that are relevant to the content of this article.

References

Bahmani, D. S., Hatzinger, M., Gerber, M., Lemola, S., Clough, P. J., Perren, S., Von Klitzing, K., Von Wyl, A., Holsboer-Trachsler, E., & Brand, S. (2016). The origins of mental toughness—prosocial behavior and low internalizing and externalizing problems at age 5 predict higher mental toughness scores at age 14. *Frontiers in Psychology, 7*, 1221. https://doi.org/10.3389/fpsyg.2016.01221

Balygina, E. A., & Ermolova, T. V. (2018). The issue of translating English psychological terms into. *Journal of Modern Foreign Psychology, 7*(1), 85–95. https://doi.org/10.17759/jmfp.2018070110

Behr, D., & Sha, M. (2018). Introduction: Translation of questionnaires in cross-national and cross-cultural research. *Translation & Interpreting: The International Journal of Translation and Interpreting Research, 10*(2), 1–4. https://doi.org/10.3316/informit.865009815260478

Bentler, P. M. (2007). On tests and indices for evaluating structural models. *Personality and Individual Differences, 42*(5), 825–829. https://doi.org/10.1016/j.paid.2006.09.024

Burton, L., Delvecchio, E., Germani, A., & Mazzeschi, C. (2021). Individualism/collectivism and personality in Italian and American groups. *Current Psychology, 40*, 29–34. https://doi.org/10.1007/s12144-019-00584-4

Chen, F., Curran, P. J., Bollen, K. A., Kirby, J. B., & Paxton, P. M. (2008). An empirical evaluation of the use of fixed cutoff points in RMSEA test statistics in structural equation models. *Sociological Methods and Research, 36*(4), 462–494. https://doi.org/10.1177/0049124108314720

Clough, P., Earle, K., & Sewell, D. (2002). Mental toughness: The concept and its measurement. In I. Cockerill (Ed.), *Solutions in Sport Psychology* (pp. 32–43). Thomson.

Connaughton, D., Wadey, R., Hanton, S., & Jones, G. (2008a). The development and maintenance of mental toughness: Perceptions of elite performers. *Journal of Sport Sciences, 26*(1), 83–95. https://doi.org/10.1080/02640410701310958

Connaughton, D., Hanton, S., Jones, G., & Wadey, R. (2008b). Mental toughness research: Key issues in this area. *International Journal of Sport Psychology, 39*(3), 192–204.

Cowden, R. G. (2018). Mental toughness inventory: Factorial validity and ethnic group measurement equivalence in competitive tennis. *Current Psychology, 39*, 736–741. https://doi.org/10.1007/s12144-018-9798-6

Crust, L., & Clough, P. J. (2005). Relationship between mental toughness and physical endurance. *Perceptual and Motor Skills, 100*(1), 192–194. https://doi.org/10.2466/pms.100.1.192-194

Crust, L., & Clough, P. J. (2011). Developing mental toughness: From research to practice. *Journal of Sport Psychology in Action, 2*(1), 21–32. https://doi.org/10.1080/21520704.2011.563436

Crust, L., & Keegan, R. (2010). Mental toughness and attitudes to risk-taking. *Personality and Individual Differences, 49*(3), 164–168. https://doi.org/10.1016/j.paid.2010.03.026

Dagnall, N., Drinkwater, K., Denovan, A., Parker, A., & Rowley, K. (2016). Misperception of chance, conjunction, framing effects and belief in the paranormal: A further evaluation. *Applied Cognitive Psychology, 30*(3), 409–419. https://doi.org/10.1002/acp.3217

Dagnall, N., Denovan, A., Papageorgiou, K. A., Clough, P. J., Parker, A., & Drinkwater, K. G. (2019). Psychometric assessment of shortened Mental Toughness Questionnaires (MTQ): Factor structure of the MTQ-18 and the MTQ-10. *Frontiers in Psychology, 10*(1933). https://doi.org/10.3389/fpsyg.2019.01933

Dagnall, N., Drinkwater, K. G., Denovan, A., & Walsh, R. S. (2021). The potential benefits of non-skills training (mental toughness) for elite athletes: Coping with the negative psychological effects of the COVID-19 pandemic. *Frontiers in Sports and Active Living, 3*, 581431. https://doi.org/10.3389/fspor.2021.581431

Denovan, A., Dagnall, N., Artamonova, E., & Papageorgiou, K. A. (2021). Dark Triad traits, learning styles, and symptoms of depression: Assessing the contribution of mental toughness longitudinally. *Learning and Individual Differences, 91*, 102053. https://doi.org/10.1016/j.lindif.2021.102053

Denovan, A., Dagnall, N., & Drinkwater, K. (2023). Examining what mental toughness, ego resiliency, self-efficacy, and grit measure: An exploratory structural equation modelling bifactor approach. *Current Psychology, 42*(26), 22148–22163. https://doi.org/10.1007/s12144-022-03314-5

Diener, E., Emmons, R. A., Larsen, R. J., & Griffin, S. (1985). The satisfaction with Life Scale. *Journal of Personality Assessment, 49*(1), 71–75. https://doi.org/10.1207/s15327752jpa4901_13

DiStefano, C., & Motl, R. W. (2006). Further investigating method effects associated with negatively worded items on self-report

surveys. *Structural Equation Modeling, 13*(2), 440–464. https://doi.org/10.1207/s15328007sem1303_6

Drinkwater, K., Dagnall, N., Denovan, A., & Parker, A. (2019). The moderating effect of mental toughness: Perception of risk and belief in the paranormal. *Psychological Reports, 122*(1), 268–287. https://doi.org/10.1177/0033294118756600

Earle, K. (2006). *The Construct of Mental Toughness: A Psychometric and Experimental Analysis.* Unpublished Ph.D. thesis, University of Hull.

Field, A., & Miles, J. (2010). *Discovering statistics using SAS*. Sage.

Galanakis, M., Lakioti, A., Pezirkianidis, C., Karakasidou, E., & Stalikas, A. (2017). Reliability and validity of the satisfaction with Life Scale (SWLS) in a Greek sample. *International Journal of Humanities and Social Studies, 5*(2), 120–127.

Gerber, M., Kalak, N., Lemola, S., Clough, P. J., Perry, J. L., Pühse, U., Elliot, C., Holsboer-Trachsler, E., & Brand, S. (2013). Are adolescents with high mental toughness levels more resilient against stress? *Stress and Health, 29*(2), 164–171. https://doi.org/10.1002/smi.2447

Gerber, M., Best, S., Meerstetter, F., Walter, M., Ludyga, S., Brand, S., Bianchi, R., Madigan, D. J., Isoard-Gautheur, S., & Gustafsson, H. (2018). Effects of stress and mental toughness on burnout and depressive symptoms: A prospective study with young elite athletes. *Journal of Science and Medicine in Sport, 21*(12), 1200–1205. https://doi.org/10.1016/j.jsams.2018.05.018

Gignac, G. E., & Szodorai, E. T. (2016). Effect size guidelines for individual differences researchers. *Personality and Individual Differences, 102*, 74–78. https://doi.org/10.1016/j.paid.2016.06.069

Goodwin, R., & Plaza, S. H. (2000). Perceived and received social support in two cultures: Collectivism and support among British and Spanish students. *Journal of Social and Personal Relationships, 17*(2), 282–291. https://doi.org/10.1177/0265407500172007

Grimm, K. J., & Widaman, K. F. (2012). Construct validity. In H. Cooper, P. M. Camic, D. L. Long, A. T. Panter, D. Rindskopf, & K. J. Sher (Eds.), *APA handbook of research methods in psychology, Vol. 1. Foundations, planning, measures, and psychometrics* (pp. 621–642). American Psychological Association. https://doi.org/10.1037/13619-033

Gucciardi, D. F. (2017). Mental toughness: Progress and prospects. *Current Opinion in Psychology, 16*, 17–23. https://doi.org/10.1016/j.copsyc.2017.03.010

Gucciardi, D., Hanton, S., & Mallett, C. (2013). Progressing measurement in mental toughness: A response to Clough, Earle, Perry and Crust. *Sport Exercise and Performance Psychology, 2*(3), 157–172. https://doi.org/10.1037/spy0000002

Heim, E., Scholten, S., Maercker, A., Xiu, D., Cai, D., Gao, Z. H., Lu, S., Sang, Z. Q., Wei, J., Kochetkov, Y., & Margraf, J. (2017). Students' value orientations in contemporary China: Analysis of measurement invariance and latent mean differences in comparison with students from Germany and Russia. *Journal of Cross-Cultural Psychology, 48*(4), 511–531. https://doi.org/10.1177/0022022117696800

Hofstede, G. (2001). Culture's recent consequences: Using dimension scores in theory and research. *International Journal of Cross-Cultural Management, 1*(1), 11–17. https://doi.org/10.1177/147059580111002

Hong, S., Malik, M. L., & Lee, M. K. (2003). Testing configural, metric, scalar, and latent mean invariance across genders in sociotropy and autonomy using a non-western sample. *Educational and Psychological Measurement, 63*(4), 636–654. https://doi.org/10.1177/0013164403251332

Horsburgh, V. A., Schermer, J. A., Veselka, L., & Vernon, P. A. (2009). A behavioural genetic study of mental toughness and personality. *Personality and Individual Differences, 46*(2), 100–105. https://doi.org/10.1016/j.paid.2008.09.009

International Test Commission (2017). *The ITC guidelines for translating and adapting tests*. http://www.intestcom.org/

Jones, D. N., & Paulhus, D. L. (2014). Introducing the short dark triad (SD3) a brief measure of dark personality traits. *Assessment, 21*(1), 28–41. https://doi.org/10.1177/1073191113514105

Kam, C. C. S., & Sun, S. (2020). Method factor due to the use of reverse-keyed items: Is it simply a response style artifact? *Current Psychology, 41*, 1204–1212. https://doi.org/10.1007/s12144-020-00645-z

Kawabata, M., Pavey, T. G., & Coulter, T. J. (2021). Evolving the validity of a mental toughness measure: Refined versions of the mental toughness questionnaire-48. *Stress and Health, 37*(2), 378–391. https://doi.org/10.1002/smi.3004

Kobasa, S. C. (1979). Stressful life events, personality, and health: An inquiry into hardiness. *Journal of Personality and Social Psychology, 37*(1), 1–11. https://doi.org/10.1037/0022-3514.37.1.1

Kuan, G., & Roy, J. (2007). Goal profiles, mental toughness and its influence on performance outcomes among Wushu athletes. *Journal of Sports Science & Medicine, 6*(CSSI-2), 28–33.

Levy, A. R., Polman, R. C., Clough, P. J., Marchant, D. C., & Earle, K. (2006). Mental toughness as a determinant of beliefs, pain, and adherence in sport injury rehabilitation. *Journal of Sport Rehabilitation, 15*(3), 245–254. https://doi.org/10.1123/jsr.15.3.245

Lima, T. J. S. D., & Souza, L. E. C. D. (2019). Rosenberg self-esteem scale: Method effect and gender invariance. *Psico-USF, 24*(3), 517–528. https://doi.org/10.1590/1413-82712019240309

Linacre, J. M. (2012). *Dimensionality: Contrasts & variances*. Retrieved from http://www.winsteps.com/winman/index.htm?principalcomponents.htm

Lindwall, M., Barkoukis, V., Grano, C., Lucidi, F., Raudsepp, L., Liukkonen, J., & Thøgersen-Ntoumani, C. (2012). Method effects: The problem with negatively versus positively keyed items. *Journal of Personality Assessment, 94*(2), 196–204. https://doi.org/10.1080/00223891.2011.645936

Loehr, J. E. (1982). *Athletic excellence: Mental toughness training for sports*. Forum Publishing Company.

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

Loehr, J. E. (1995). *The new toughness training for sports: Mental emotional physical conditioning from one of the world's premier sports psychologists*. Penguin.

Mack, M. G., & Ragan, B. G. (2008). Development of the mental, emotional, and bodily toughness inventory in collegiate athletes and nonathletes. *Journal of Athletic Training, 43*(2), 125–132. https://doi.org/10.4085/1062-6050-43.2.125

Marchant, D. C., Polman, R. C., Clough, P. J., Jackson, J. G., Levy, A. R., & Nicholls, A. R. (2009). Mental toughness: Managerial and age differences. *Journal of Managerial Psychology, 24*(5), 428–437. https://doi.org/10.1108/02683940910959753

Marsh, H. W. (1996). Positive and negative global self-esteem: A substantively meaningful distinction or artifactors? *Journal of Personality and Social Psychology, 70*(4), 810–819. https://doi.org/10.1037/0022-3514.70.4.810

Marsh, H. W., Vallerand, R. J., Lafrenière, M. A. K., Parker, P., Morin, A. J., Carbonneau, N., Jowett, S., Bureau, J. S., Fernet, C., Guay, F., Salah Abduljabbar, A., & Paquet, Y. (2013). Passion: Does one scale fit all? Construct validity of two-factor passion scale and psychometric invariance over different activities and languages. *Psychological Assessment, 25*(3), 796–809. https://doi.org/10.1037/a0032573

Mojtahedi, D., Dagnall, N., Denovan, A., Clough, P., Hull, S., Canning, D., Lilley, C., & A Papageorgiou, K. (2020). The Relationship between mental toughness, job loss, and mental health issues during the COVID-19 pandemic. *Frontiers in Psychiatry, 11*. https://doi.org/10.3389/fpsyt.2020.607246

Mojtahedi, D., Dagnall, N., Denovan, A., Clough, P., Dewhurst, S., Hillier, M., Papageorgiou, K. A., & Perry, J. (2023). Competition anxiety in combat sports and the importance of mental toughness. *Behavioral Sciences, 13*(9), 713. https://doi.org/10.3390/bs13090713

Nicholls, A. R., Polman, R. C., Levy, A. R., & Backhouse, S. H. (2008). Mental toughness, optimism, pessimism, and coping among athletes. *Personality and Individual Differences, 44*(5), 1182–1192. https://doi.org/10.1016/j.paid.2007.11.011

Onley, M., Veselka, L., Schermer, J. A., & Vernon, P. A. (2013). Survival of the scheming: A genetically informed link between the dark triad and mental toughness. *Twin Research and Human Genetics, 16*(6), 1087–1095. https://doi.org/10.1017/thg.2013.66

Papageorgiou, K. A., Wong, B., & Clough, P. J. (2017). Beyond good and evil: Exploring the mediating role of mental toughness on the Dark Triad of personality traits. *Personality and Individual Differences, 119*, 19–23. https://doi.org/10.1016/j.paid.2017.06.031

Papageorgiou, K. A., Malanchini, M., Denovan, A., Clough, P. J., Shakeshaft, N., Schofield, K., & Kovas, Y. (2018). Longitudinal associations between narcissism, mental toughness and school achievement. *Personality and Individual Differences, 131*, 105–110. https://doi.org/10.1016/j.paid.2018.04.024

Papageorgiou, K. A., Denovan, A., & Dagnall, N. (2019). The positive effect of narcissism on depressive symptoms through mental toughness: Narcissism may be a dark trait but it does help with seeing the world less grey. *European Psychiatry, 55*, 74–79. https://doi.org/10.1016/j.eurpsy.2018.10.002

Papageorgiou, K. A., Denovan, A., Dagnall, N., Artamonova, E., Gianniou, F. M., Papageorgiou, S., Plouffe, R. A., Kowalski, C. M., Saklofske, D. H., Kyriazos, T., Stalikas, A., & Costantini, G. (2023). Grandiose narcissism indirectly associates with lower psychopathology across five countries. *Journal of Psychiatric Research, 167*, 78–85. https://doi.org/10.1016/j.jpsychires.2023.10.003

Pavot, W., & Diener, E. (2008). The satisfaction with Life Scale and the emerging construct of life satisfaction. *The Journal of Positive Psychology, 3*(2), 137–152. https://doi.org/10.1080/17439760701756946

Perry, J. L., Nicholls, A. R., Clough, P. J., & Crust, L. (2015). Assessing model fit: Caveats and recommendations for confirmatory factor analysis and exploratory structural equation modeling. *Measurement in Physical Education and Exercise Science, 19*(1), 12–21. https://doi.org/10.1080/1091367X.2014.952370

Perry, J. L., Strycharczyk, D., Dagnall, N., Denovan, A., Papageorgiou, K. A., & Clough, P. J. (2021). Dimensionality of the Mental Toughness Questionnaire (MTQ48). *Frontiers in Psychology, 12*, 654836. https://doi.org/10.3389/fpsyg.2021.654836

Podsakoff, P. M., MacKenzie, S. B., Lee, J. Y., & Podsakoff, N. P. (2003). Common method biases in behavioral research: A critical review of the literature and recommended remedies. *Journal of Applied Psychology, 88*(5), 879–903. https://doi.org/10.1037/0021-9010.88.5.879

Podsakoff, P. M., MacKenzie, S. B., & Podsakoff, N. P. (2012). Sources of method bias in social science research and recommendations on how to control it. *Annual Review of Psychology, 63*, 539–569. https://doi.org/10.1146/annurev-psych-120710-100452

Rogers, R., Jackson, R. L., Sewell, K. W., & Salekin, K. L. (2005). Detection strategies for malingering: A confirmatory factor analysis of the SIRS. *Criminal Justice and Behavior, 32*(5), 511–525. https://doi.org/10.1177/0093854805278412

Rutkowski, L., & Svetina, D. (2014). Assessing the hypothesis of measurement invariance in the context of large-scale international surveys. *Educational and Psychological Measurement, 74*(1), 31–57. https://doi.org/10.1177/0013164413498257

Sabouri, S., Gerber, M., Sadeghi Bahmani, D., Lemola, S., Clough, P. J., Kalak, N., Shamsi, M., Holsboer-Trachsler, E., & Brand, S. (2016). Examining Dark Triad traits in relation to mental toughness and physical activity in young adults. *Neuropsychiatric Disease and Treatment*, 229–235. https://doi.org/10.2147/NDT.S97267

Schumacker, R. E., & Lomax, R. G. (2015). *A beginner's guide to structural equation modeling* (4th ed.). Routledge.

Sola, V., Sakiyama, H. M. T., Padin, R., Canfield, M. D. F., Bortolon, M., Laranjeira, C. B., R., & Mitsuhiro, S. S. (2018). Measuring stress, coping, strain and hopefulness of Brazilian family members of substance misusers: Factor structure of a set of measures. *Journal of Substance Use, 24*(2), 130–139. https://doi.org/10.1080/14659891.2018.1523963

Somma, A., Paulhus, D. L., Borroni, S., & Fossati, A. (2019). Evaluating the psychometric properties of the short Dark Triad (SD3) in Italian adults and adolescents. *European Journal of Psychological Assessment, 36*(1), 185–195. https://doi.org/10.1027/1015-5759/a000499

Son, J. (2018). Back translation as a documentation tool. *Translation & Interpreting, 10*(2), 89–100. https://doi.org/10.12807/ti.110202.2018.a07

Sonderen, E. V., Sanderman, R., & Coyne, J. C. (2013). Ineffectiveness of reverse wording of questionnaire items: Let's learn from cows in the rain. *PloS ONE, 8*(7), e68967. https://doi.org/10.1371/journal.pone.0068967

Stamatis, A., Morgan, G. B., Papadakis, Z., Mougios, V., Bogdanis, G., & Spinou, A. (2021). Cross-cultural invariance of the mental toughness index among American and Greek athletes. *Current Psychology, 40*, 5793–5800. https://doi.org/10.1007/s12144-019-00532-2

Strycharczyk, D., & Clough, P. J. (2014). *Developing mental toughness in young people: Approaches to achievement, well(-)being, employability, and positive Behaviour*. Karnac Books.

Szabo, Z. P., Kun, A., Balogh, B. E., Simon, E., & Csike, T. (2022). Dark and Strong?! The associations between dark personality traits, mental toughness and resilience in Hungarian student, employee, leader, and military samples. *Personality and Individual Differences, 186*, 111339. https://doi.org/10.1016/j.paid.2021.111339

Tabachnick, B. G., & Fidell, L. S. (2013). *Using multivariate statistics* (6th ed.). Allyn and Bacon.

Triandis, H. C., Chan, D. K. S., Bhawuk, D. P., Iwao, S., & Sinha, J. B. (1995). Multimethod probes of allocentrism and idiocentrism. *International Journal of Psychology, 30*(4), 461–480. https://doi.org/10.1080/00207599508246580

Wheatley, C., Batey, M., Denovan, A., & Dagnall, N. (2023). Mental toughness in the Football Association women's Super League: Relationships with playing experience, perceptions of club infrastructure, support mechanisms and self-esteem. *PloS One, 18*(5), e0285594. https://doi.org/10.1371/journal.pone.0285594