

Zur Publikation in der Zeitschrift *Diagnostica* angenommenes Manuskript  
(<https://doi.org/10.1026/0012-1924/a000192> © 2017 Hogrefe-Verlag). Diese Artikelfassung  
entspricht nicht vollständig dem in der Zeitschrift veröffentlichten Artikel. Dies ist nicht die  
Originalversion des Artikels und kann daher nicht zur Zitierung herangezogen werden.

Psychometrische Eigenschaften einer deutschsprachigen Übersetzung  
des Mental Toughness Inventory (MTI-D)

Psychometric Properties of a German Translation  
of the Mental Toughness Inventory (MTI-D)

Jürg Schmid, Silvan Steiner und Manuela Rechsteiner Rensch  
Universität Bern

Cory Middleton  
Talent Development Academy, Sydney, Australien

Roland Seiler  
Universität Bern

Korrespondenzadresse: Jürg Schmid, Institut für Sportwissenschaft, Universität Bern,  
Bremgartenstrasse 145, CH-3012 Bern. E-Mail: [juerg.schmid@ispw.unibe.ch](mailto:juerg.schmid@ispw.unibe.ch)

**Kurztitel:** Gütekriterien des MTI-D

**Schlüsselwörter:** mentale Stärke; deutschsprachige Version des Mental Toughness Inventory (MTI); Gütekriterien; exploratorische Strukturgleichungsmodelle (ESEM); Bifaktor-Modell

**Key words:** mental toughness; German version of the Mental Toughness Inventory (MTI); psychometric properties; exploratory structural equation modeling (ESEM); bifactor model

**Zusammenfassung:** Das Mental Toughness Inventory (MTI; Middleton, Martin & Marsh, 2011) ist ein Selbstbeurteilungsinstrument zur Erfassung von 12 Facetten mentaler Stärke im Sport oder anderen leistungsbezogenen Kontexten. Ins Deutsche übersetzt, wurde das MTI-D an einer deutschsprachigen Stichprobe ( $N = 1\,122$ ), davon  $n = 914$  Athletinnen und Athleten und  $n = 208$  Studierende der Sportwissenschaft aus 41 Sportarten hinsichtlich Faktorstruktur, Messinvarianz, Reliabilität, Konstrukt und Kriteriumsvalidität untersucht. Entsprechend der theoretischen Konzeption ergaben Bifaktorenanalysen für das MTI-D eine Struktur mit einem gemeinsamen Faktor und 12 spezifischen Faktoren. Die Reliabilität ( $\omega$ ) der Gesamtskala und der Subskalen lag zwischen .81 und .98, während die Werte für omega hierarchical subscale tiefer lagen ( $.07 \leq \omega_{HS} \leq .77$ ). Der globale und die spezifischen Faktoren korrelierten hypothesenkonform mit dem Engagement und Leistungsniveau im Sport sowie dem Test of Performance Strategies (TOPS-D; Schmid, Birrer, Kaiser & Seiler, 2010) zur Erfassung mentaler Strategien im Sport. Zwar könnten einzelne Faktoren optimiert werden, doch es liegen Hinweise darauf vor, dass mit dem MTI-D mentale Stärke reliabel und valide erfasst werden kann.

**Abstract:** The Mental Toughness Inventory (MTI; Middleton, Martin & Marsh, 2011) is a self-description questionnaire measuring 12 facets of mental toughness in sports or other performance-oriented contexts. Translated into German, the MTI-D was analysed with respect to its factorial structure, measurement invariance, reliability, construct validity, and criterion validity using a German-speaking sample of  $N = 1\,122$  individuals including  $n = 914$  athletes and  $n = 208$  sport science students involved in 41 different sports. In line with the theoretical conception of the MTI and using a (bifactor) exploratory and confirmatory structural equation modelling framework, the analyses of the MTI-D yielded a factor structure with one general factor and 12 domain-specific factors. The results supported

measurement invariance of the MTI-D across gender and age. Omega estimates of composite reliability ( $\omega$ ) for the global factor and the domain-specific factors ranged between .81 and .91, while omega hierarchical subscale was much lower ( $.07 \leq \omega_{HS} \leq .77$ ). The MTI-D factors correlated in predictable ways with degree of sport involvement, level of achievement in sport, and the Test of Performance Strategies (TOPS-D; Schmid, Birrer, Kaiser & Seiler, 2010), an instrument measuring mental strategies in sports. Despite that some specific factors await further refinement, there is preliminary evidence that the MTI-D can be used as a reliable and valid questionnaire for assessing mental toughness.

Seit Loehrs Beitrag aus dem Jahre 1986 wird mentale Stärke – zusammen mit dem athletischen Potential – konsensual als wichtiger Faktor für das Erbringen guter Leistungen im Sport angesehen. Weniger Einigkeit besteht hinsichtlich der Fragen, welche Aspekte mentale Stärke konstituieren und wie sie erlangt wird (Richards, 2011). Gemäß Sheard (2012) sollte mentale Stärke als Persönlichkeitsstil und kognitive Haltung konzeptualisiert sein, die durch Kultur, gesellschaftliche Rahmenbedingungen, Erfahrung, Training und vermutlich auch genetische Aspekte (Horsburgh, Schermer, Veselka & Vernon, 2009) bestimmt wird. Im Detail regt Sheard (2012) an, dass mentale Stärke theoretisch in die positive Psychologie eingebettet wird, dass ihr eine Moderatorfunktion im reaktiven Umgang mit wettkampfinduziertem Stress zukommt und dass sie proaktive Komponenten berücksichtigt, da sich die Wirkung mentaler Stärke auf die sportliche Leistung schon vor der Auseinandersetzung mit aversiven Situationen entfaltet.

Zur Erfassung mentaler Stärke sind im englischen Sprachraum mehrere Instrumente verfügbar (im Überblick: Gucciardi, Mallet, Hanrahan & Gordon, 2011; Sheard, 2012), namentlich der Mental Toughness Questionnaire (MTQ48; Clough, Earle & Sewell, 2002), der Sport Mental Toughness Questionnaire (SMTQ; Sheard, Golby & van Wersch, 2009) und das Mental Toughness Inventory (MTI; Middleton, Marsh, Martin, Richards & Perry, 2005a, 2005b). Als wir mit unserer Studie begannen, war im deutschen Sprachraum kein Instrument verfügbar. Gerber et al. haben jedoch 2013 den MTQ48 übersetzt.

Um diese Lücke zu schließen, erachteten wir es als sinnvoll, am MTI anzuknüpfen. Dies erschien interessant wegen des kombinierten Top-down- und Bottom-up-Ansatzes (s.

unten), auf dem seine Entwicklung basiert, und der vorgelegten psychometrischen Analysen, speziell zur faktoriellen Validität. Zudem ist das MTI anders als der SMTQ auch in leistungsbezogenen Kontexten außerhalb des Sports einsetzbar und erlaubt eine differenzierte Erfassung der mentalen Stärke. Dies kann der MTQ48, der bislang zur Erfassung von mentaler Stärke am häufigsten eingesetzt worden ist, auch für sich beanspruchen. Er ist aber heftig kritisiert worden (Connaughton & Hanton, 2010; Gucciardi et al., 2011), weil Dimensionsanalysen zu inkonsistenten oder sonst unbefriedigenden Befunden führten. Seine Faktorstruktur bleibt damit unklar (für Befunde zur deutschsprachigen Version s. Gerber et al., 2013). Gucciardi, Hanton und Malletts (2012) Bilanz ist, dass die theoretische Konzeption des MTQ48 mit vier, sechs oder neun Dimensionen grundlegend zu überdenken ist.

Ziel dieses Beitrags ist es, Analysen zur psychometrischen Qualität des MTI-D vorzustellen, d.h. unserer Übersetzung des MTI ins Deutsche, die wir Vallerands (1989) Verfahren folgend vorgenommen haben. Beispielitems des MTI-D sind im Elektronischen Supplement (ESM) 1 verfügbar, die vollständige Version ist bei Dr. C. Middleton zu beziehen.<sup>1</sup>

### *Entwicklung und psychometrische Eigenschaften des MTI*

Das MTI wurde in einem zweistufigen Verfahren unter Verknüpfung qualitativer und quantitativer Methoden entwickelt: Im ersten Schritt fragten Middleton, Marsh, Martin, Richards und Perry (2004b) 33 Elite-Trainerinnen und -Trainer sowie Elite-Athletinnen und Athleten nach ihrer subjektiven Konzeptualisierung von mentaler Stärke. Auf dieser Basis definierten Middleton, Martin und Marsh (2011, S. 94) mentale Stärke als „an unshakeable perseverance and conviction towards some goal despite pressure or adversity“.

Im zweiten Schritt wurden die subjektiven Konzeptualisierungen vor dem Hintergrund etablierter psychologischer Theorien systematisiert (Middleton et al., 2004b). Entstanden ist ein mehrdimensionales, hierarchisches Modell (s. ESM 2, Modell 0), in welchem mentale Stärke als übergeordnete globale Fähigkeit konzipiert ist, die in Orientierung und Strategie untergliedert wird. Unter Orientierung werden persönliche Charakteristika und unter Strategie Handlungen subsumiert. Orientierung umfasst die Faktoren *Selbstwirksamkeit*,

---

<sup>1</sup> Dr. Cory Middleton, Talent Development Academy, Sydney, Australia, E-Mail: cory@tdacademy.com.au, Phone +61 409 905 152.

*Selbstkonzept mentaler Stärke*, der Glaube an das eigene *Potential*, *Vertrautheit mit der Aufgabe*, *Streben nach (persönlicher) Bestleistung*, *Aufgabenvolenz* und *Zielbindung*. Strategie dagegen setzt sich zusammen aus den Faktoren *Durchhaltevermögen*, *Aufgabenbezogene Aufmerksamkeit*, *Stressminimierung*, *Positive Einstellung* und *Positive Vergleiche*. Damit berücksichtigt das MTI 12 Faktoren, die sich unter anderem an Theorien zum Stresserleben, zur Selbstwirksamkeit, zur Selbstbestimmung und zum Selbstkonzept anlehnen.

Die erste Version des MTI (Middleton, Marsh, Martin, Richards & Perry, 2004a) umfasste 67 Items, die auf einer achtstufigen Likert-Skala abgefragt wurden. Analysen an 479 Jugendlichen in sportorientierten Schulen ergaben eine gute Passung für ein konfirmatorisches Faktorenmodell erster Ordnung (s. ESM 2, Modell A) und für alle 12 Skalen eine hohe innere Konsistenz (Cronbachs  $\alpha$ ) von .87 bis .95.

In der aktuellen Version des MTI (Middleton et al., 2005a) wurde die Zahl der Items auf drei pro Skala, insgesamt 36, reduziert. Zudem wurde die Berechnung eines Gesamtscores vorgeschlagen. Untersuchungen an 438 angehenden und 392 etablierten Elite-Sportlerinnen und -Sportlern auch für die gekürzte Version bestätigten die hohen inneren Konsistenzen ( $.82 \leq \alpha \leq .94$ ). Analysen auf der Basis des Faktorenmodells erster Ordnung (Modell A) ergaben in beiden Stichproben erneut einen guten Modellfit (angehende Elite  $\chi^2(528, N = 438) = 1\,275.15$ , TLI = .99, RMSEA = .057; Elite:  $\chi^2(528, N = 392) = 926.75$ , TLI = .99, RMSEA = .044).

Für die ursprüngliche MTI-Version (67 Items) wurden mittlere Korrelationen mit psychologischen Konstrukten wie dem physischen Selbstkonzept und Flow ermittelt und für die 36-Item-Version zusätzlich Zusammenhänge mit Kontrollüberzeugungen und Personal Effectiveness gefunden. Middleton et al. (2004a, 2005a, 2011) zogen das Fazit, dass sich die Skalen gut bewähren. Angaben zur Kriteriumsvalidität des MTI finden sich bislang erst in der Arbeit von Middleton et al. (2005b). In ihrer Gegenüberstellung angehender und etablierter Elite-Sportlerinnen und -Sportler ergaben sich in den Skalen *Selbstwirksamkeit*, *Vertrautheit mit der Aufgabe*, *Stressminimierung*, *Selbstkonzept mentaler Stärke*, *Positive Einstellung* und *Aufgabenbezogene Aufmerksamkeit* jedoch relativ geringe Ausprägungen für die etablierten Elite-Sportlerinnen und -Sportler. Middleton et al. (2005b) vermuteten, dass diese etablierten Athletinnen und Athleten eine selektive Stichprobe darstellen, die ihre Fähigkeiten im Kontext relativ starker Sportlerinnen und Sportler einschätzten.

Die Abklärungen zur psychometrischen Qualität des MTI sind so weit fortgeschritten, dass Middleton et al. (2011) das Instrument zum Einsatz in Praxis und Forschung vorschlagen. Aus unserer Warte ist jedoch kritisch auf die hohen Faktorkorrelationen in den beiden Validierungsstichproben hinzuweisen, die im besseren Fall durchschnittlich  $.45$  ( $.13 \leq r \leq .69$ ) und im schlechteren durchschnittlich  $.58$  ( $.32 \leq r \leq .81$ ) betragen (Middleton et al., 2005a). Wenn die einzelnen Faktoren wenig eigenständige Aspekte von mentaler Stärke erfassen, schränkt dies nicht nur die Chance ein, Hinweise auf die diskriminante Validität der Faktoren zu finden, sondern möglicherweise auch den diagnostischen Nutzen des Instruments.

Die Gruppe um Middleton (Middleton et al., 2005a) hat es damals beim erwähnten CFA-Modell bewenden lassen (müssen), als sie die Faktorstruktur des MTI analysierten. Ihre Analyse gründete auf dem restriktiven *independent cluster model* (ICM; McDonald, 1999), das eine perfekte Einfachstruktur der Faktorladungsmatrix unterstellt. In den letzten Jahren sind jedoch flexiblere Ansätze vorgeschlagen und mit Erfolg zur Analyse mehrdimensionaler Konstrukte eingesetzt worden, namentlich exploratorische Strukturgleichungsmodelle (ESEM; Marsh et al., 2009), Bifaktormodelle (Reise, 2012) und deren Kombination (Morin, Arens & Marsh, 2016).

Der ESEM-Ansatz verbindet die Vorteile der exploratorischen Faktorenanalyse (EFA) mit jenen der konfirmatorischen (CFA), nämlich die Berechnung exploratorischer Faktorenmodelle unter Angabe der für die CFA üblichen Teststatistiken (Asparouhov & Muthén, 2009). Speziell angezeigt ist der ESEM-Ansatz aber, weil er explizit Nebenladungen zulässt und nicht Modelle mit perfekter Einfachstruktur testet, wie dies bei der CFA meist der Fall ist (ICM-Ansatz). Im Ausschluss solcher Nebenladungen sehen Marsh et al. (2009) den Grund für schlecht passende CFA-Modelle und überhöhte Zusammenhänge zwischen Faktoren.

Bifaktormodelle erklären die manifesten Variablen durch einen globalen oder gemeinsamen Faktor (G) sowie mehrere spezifische Faktoren (S), die als voneinander unabhängig angenommen werden (Reise, 2012). Dies steht im Gegensatz zu den Faktorenmodellen höherer Ordnung, in denen der Globalfaktor die spezifischen Faktoren direkt und die einzelnen Items indirekt beeinflusst. Dieses Bifaktormodell ist für das MTI auch theoretisch interessant, da Middleton et al. (2011) bei der Entwicklung einen Globalfaktor vor Augen hatten, ihn angesichts der empirischen Befunde aber zurückstellen

mussten.

Insgesamt handelt es sich beim MTI um ein Instrument, das auf einem originellen Modell (Gerber, 2011) gründet. Es liegen auch schon vielversprechende Analysen zu dessen testpsychologischen Güte vor, wobei noch keine Studie publiziert ist, welche die hierarchische und multidimensionale Struktur geprüft hat.

### *Zielsetzung und Hypothesen*

Ziel der vorliegenden Studie ist, die psychometrische Qualität des MTI-D unter vier Aspekten zu analysieren:

Erstens betrachten wir dessen faktorielle Validität und untersuchen, inwieweit sich die von Middleton et al. (2005a) überprüfte Struktur (Modell A, ICM-CFA) bestätigen lässt. Zudem prüfen wir die Datenpassung dreier weiterer Modelle (s. ESM 2): Modell B ist ein von Middleton et al. (2011) postuliertes, aber nicht getestetes hierarchisches CFA-Modell (HCFA) mit 12 Faktoren erster und einem Faktor zweiter Ordnung. Bei Modell C geht es um ein exploratorisches Faktorenmodell erster Ordnung (ESEM), das anders als Modell A keine perfekte Einfachstruktur der Faktorladungsmatrix unterstellt. Modell D schließlich ist ein exploratorisches Bifaktormodell (BESEM), das ebenfalls Nebenladungen zulässt. Für das beste Modell wird anhand des Geschlechts, des Alters und der Zugehörigkeit zur Gruppe der Vereinssportlerinnen und -sportler bzw. der Sportstudierenden die konfigurale, metrische und skalare Messinvarianz geprüft.

Zweitens untersuchen wir die Reliabilität des MTI-D. Dazu wird die Methode der Konsistenzanalyse sowie der Testwiederholung verwendet, wobei aus Platzgründen alle Ausführungen zur Retest-Reliabilität im Elektronischen Supplement 2 verfügbar sind.

Drittens ermitteln wir die Beziehung zwischen dem MTI-D und dem Test of Performance Strategies (TOPS-D; Schmid, Birrer, Kaiser & Seiler, 2010). Beide Instrumente beanspruchen, psychische Voraussetzungen optimaler Leistungen im Sport zu erfassen. Allgemein wird erwartet, dass sich diese Gemeinsamkeit in einer moderaten Korrelation der MTI-D- und TOPS-D-Skalen niederschlägt. Plausibilitätsüberlegungen und Ausführungen von Middleton et al. (2004) zufolge formulieren wir vier spezifische Hypothesen:

(a) Hohe Ausprägungen auf den MTI-D-Skalen *Stressminimierung* und *Positive Einstellung* gehen mit geringen Ausprägungen auf der TOPS-D-Skala *Negative Kognitionen* einher: Eine wirksame Strategie zur Stressminimierung und eine ausgeprägte Tendenz, auch

negativen Dingen etwas Positives abzugewinnen, führen zu weniger negativen Kognitionen.

(b) Derselben Überlegung folgend erwarten wir, dass hohe Werte auf den beiden MTI-D-Skalen *Stressminimierung* und *Positive Einstellung* mit hohen Werten auf der TOPS-D-Skala *Emotionale Kontrolle* einhergehen.

(c) Für die MTI-D-Skalen *Selbstkonzept mentaler Stärke* und *Selbstwirksamkeit* wird ein negativer Zusammenhang mit der TOPS-D-Skala *Negative Kognitionen* vermutet, und zwar aufgrund der gegenseitigen Abhängigkeit von mentalem Selbstkonzept und Kognitionen (Shavelson, Hubner & Stanton, 1976) und aufgrund der kognitiven Natur von Selbstwirksamkeitserwartungen (Bandura, 1997).

(d) Zudem erwarten wir einen positiven Zusammenhang zwischen *Aufgabenbezogene Aufmerksamkeit* (MTI-D) und *Aufmerksamkeitskontrolle* (TOPS-D), da sich beide Skalen auf die Aufmerksamkeitslenkung bei sportlichen Aufgaben beziehen.

Viertens wird die Kriteriumsvalidität des MTI-D überprüft. Wir vermuten, dass mentale Stärke während sportlicher Aktivität entwickelt und erworben wird und dass solche Sozialisationsprozesse in dem Maße einflussreich sind, in dem über lange Zeit, oft und intensiv Wettkampfsport betrieben wird (Sheard, 2012). Auch Selektionsprozesse sind anzunehmen, weil mental (und athletisch) weniger starke Sportlerinnen und -sportler in ihren Leistungen stagnieren, sich eher im Breiten- als im Leistungssport betätigen oder den Sport ganz aufgeben. Wir erwarten also, dass der MTI-D mit dem Engagement und Leistungsniveau im Sport korrelieren.

## *Methode*

### *Stichprobe*

Für die Bearbeitung der aufgeworfenen Fragestellungen wurden in unterschiedlichen Teilstichproben je spezifische Daten gesammelt. Untersucht wurden 208 Studierende der Sportwissenschaft der Universität Bern im ersten Semester (Vollerhebung zweier Jahreshkohorten zu 93 bzw. 115 Personen; Alter:  $M = 21.04$  Jahre,  $SD = 1.96$ ; 50 % Frauen) sowie 914 unterschiedlich leistungsstarke Vereinssportlerinnen und -sportler, die in 41 Sportarten aktiv waren (Schneeballverfahren; Alter:  $M = 22.83$  Jahre,  $SD = 4.13$ ; 40 % Frauen). Zur Analyse der faktoriellen Validität, Messinvarianz und inneren Konsistenz wurde die gesamte Stichprobe herangezogen ( $N = 1\,122$ ), während zur Beurteilung der Konstrukt- und Kriteriumsvalidität die zwei Jahreshkohorten der Sportstudierenden dienten.



### *Messinstrumente zur Validierung*

Ein wichtiger Ausgangspunkt der Entwicklung des MTI war das Psychological Performance Inventory (PPI; Loehr, 1986), das noch bis in die 00er-Jahre in der Beratungspraxis als das bedeutendste Instrument zur Erfassung von mentaler Stärke galt (Middleton et al., 2004a). Testpsychologisch betrachtet erwies sich jedoch weder die englisch- noch die deutschsprachige Version als haltbar (Middleton et al., 2004a; Schmid & Gubelmann, 1999). Da zu Beginn unserer Studie kein passables Instrument zur Erfassung mentaler Stärke in deutscher Sprache vorlag, wählten wir als Validierungsinstrument den TOPS-D (Schmid et al., 2010), der mit 16 Skalen den Einsatz mentaler Strategien im Kontext von Training und Wettkampf erfasst. Zur Reduktion der  $\alpha$ -Fehler-Kumulierung wurden kontextübergreifende Skalen gebildet, so dass letztlich neun Dimensionen psychischer Fertigkeiten und Techniken gemessen wurden. Die Skalen *Emotionale Kontrolle*, *Automatisierung*, *Aktivierung*, *Aufmerksamkeitskontrolle* und *Negative Kognitionen* erfassen das Ausmaß, in dem es Sportlerinnen und Sportlern gelingt, sich im entscheidenden Zeitpunkt in jenen Zustand zu bringen, der für das Erbringen sportlicher Leistungen optimal ist. Mit den Skalen *Selbstgespräch*, *Zielsetzung*, *Visualisierung* und *Entspannung* werden Techniken gemessen, auf die Athletinnen und Athleten zurückgreifen, um diesen optimalen Zustand zu erreichen. In der vorliegenden Studie schwankte die innere Konsistenz der Skalen bei einem Mittelwert von .79 zwischen .68 und .91.

Das Leistungsniveau im Sport wurde mit einem Item erhoben, in dem die Probanden gefragt wurden, ob sie mit der nationalen Meisterin oder dem nationalen Meister in ihrer Disziplin sportlich mithalten können. Die siebenstufige Antwortskala reichte von 1 (*Nein, auf keinen Fall*) bis 7 (*Ja, auf jeden Fall*). Das Sportengagement wurde mit zweimal drei Items abgedeckt: Die Wettkampftätigkeit (*keine, regionale, inter-/nationale Wettkämpfe*), die Anzahl wöchentlicher Trainingseinheiten und den wöchentlichen Trainingsaufwand in Stunden. Die Angaben wurden retrospektiv für den 1-Jahres-Zeitraum vor der Erhebung und den ihm vorangehenden 10-Jahres-Zeitraum erfragt. Zur Optimierung der Verteilungsform (Bortz & Schuster, 2010) wurden die Items zum Trainingsaufwand quadratwurzeltransformiert. Mit dem Ziel der Datenreduktion und Skalenbildung wurden die zweimal drei Items einer Hauptkomponentenanalyse (Promax-Rotation) unterzogen. Extrahiert wurden zwei Hauptkomponenten (Varianzklärungsanteil: 75 %;  $r = .32$ ). Es ergab sich weitestgehend eine Einfachstruktur, wobei die Items zum 1-Jahres-Zeitraum hoch

auf die eine Hauptkomponente luden ( $.73 \leq \lambda \leq .94$ ) und die Items zum 10-Jahres-Zeitraum auf die andere ( $.73 \leq \lambda \leq .93$ ). Die Hauptkomponenten werden als vergangenes und aktuelles Sportengagement interpretiert.

#### *Datenanalyse*

Fehlende Werte in den Datensätzen wurden mithilfe des Expectation-Maximation-Algorithmus einfach imputiert.

Die faktorielle Validität wurde mithilfe der Faktorenanalyse untersucht (Mplus 7.4; Muthén & Muthén, 1998-2015). Zur Beurteilung der Modellgüte wurden vier Maße und Cut-off-Werte herangezogen (Brown, 2015): CFI ( $\approx .95$ ), TLI ( $\approx .95$ ), RMSEA ( $< .06$ ) und SRMR ( $< .08$ ). Marsh et al. (2009) folgend stützten wir uns auch für ESEM-Modelle auf diese Cut-off-Werte, auch wenn noch ungeklärt ist, ob sie hierfür angemessen sind. Genestete Modelle wurden anhand der nach Satorra und Bentler (2001) skalierten  $\chi^2$ -Differenz (*TRd*) miteinander verglichen. Chen (2007) folgend wurde aber in erster Linie auf die CFI- und RMSEA-Differenz der Modelle abgestellt und die Nullhypothese bei einer Differenz  $> .010$  bzw.  $.015$  abgelehnt.

Zur Klärung der Frage nach dem passenden Modell wurde auf eine Auswertungsstrategie von Morin et al. (2016) rekurriert. Sie besteht darin, in einem ersten Schritt das ICA-CFA-Modell A und in einem zweiten das HCFA-Modell (B) sowie das ESEM-Modell (C) zu schätzen und so ein Modell mit möglichst unverzerrten Schätzungen der Faktorkorrelationen zu finden. Dieses Modell soll schließlich dem entsprechenden Bifaktormodell (BCFA oder BESEM) gegenübergestellt werden: So kann direkt geprüft werden, ob allenfalls ein G-Faktor vorliegt, der allen Items zugrunde liegt und mit den multiplen spezifischen Faktoren (S) koexistiert.

*Messinvarianz:* Nach der Bestimmung des theoretisch und empirisch besten Modells wurde geprüft, ob das MTI-D über zwei Sportlerinnen bzw. Sportler- (Vereinsportlerinnen bzw. -sportler vs. Sportstudierende), Geschlechts- (Frauen vs. Männer) und Altersgruppen (Mediansplit,  $Md = 22.0$  Jahre) äquivalent misst. Konkret wurde das MTI-D mit multiplen Gruppenvergleichen im Rahmen der Strukturgleichungsmodellierung sequenziell auf konfigurale (gleiche Faktorenstruktur), metrische (zusätzlich gleiche Faktorenladungen) und skalare (zusätzlich gleiche Itemmittelwerte) Messinvarianz getestet (Brown, 2015). Dies geschah wiederum anhand der CFI- und RMSEA-Differenz der Modelle.

*Reliabilität:* Die Reliabilitätsschätzung erfolgte über die Konsistenzanalyse, d.h.

anhand von Cronbachs  $\alpha$  sowie vier Typen von McDonalds (1999)  $\omega$ : omega ( $\omega$ ), omega subscale ( $\omega_S$ ), omega hierarchical ( $\omega_H$ ) und omega hierarchical subscale ( $\omega_{HS}$ ) (Rodriguez, Reise & Haviland, 2016).  $\omega$  und  $\omega_S$  liefern Informationen zur Reliabilität der Gesamtskala bzw. der Subskalen;  $\omega_H$  beziffert den Anteil reliabler Varianz des Globalfaktors;  $\omega_{HS}$  gibt den Anteil reliabler Varianz eines spezifischen Faktors an, nachdem die reliable Varianz des Globalfaktors auspartialisiert ist. Die Berechnung durchschnittlicher Reliabilitäten und Korrelationen wurde via Fishers Z-Werte durchgeführt (Bortz & Schuster, 2010).

Zur Beurteilung individueller Differenzen wurden Reliabilitäten ( $\alpha$ ,  $\omega$ ,  $\omega_H$ ,  $\omega_S$ ) von .70, zur Beurteilung von Gruppendifferenzen solche von .50 als noch ausreichend betrachtet, wengleich höhere Werte erwünscht sind (Lienert & Raatz, 1998). Für  $\omega_{HS}$  fehlen Konventionen zur Beurteilung. Reise, Bonifay und Haviland (2013, S. 137) schlagen „vorläufig“ und „subjektiv“ ein Minimum von .50 vor, „präferieren“ aber Werte um .75.

*Konstrukt- und Kriteriumsvalidität:* Die Analysen zur Konstrukt- und Kriteriumsvalidität gründeten auf kleinen Stichproben ( $n = 93$  bzw.  $n = 115$ ). Strukturgleichungsmodelle waren deshalb keine Option, so dass auf manifester Ebene Rangkorrelationen berechnet wurden, wobei das Interesse den Korrelationen zwischen MTI-D einerseits und TOPS-D, Sportengagement und Leistungsniveau andererseits galt. Konkret wurden herangezogen: die MTI-D-Faktorwerte, die A-priori-Skalen des TOPS-D, die Faktorwerte für das aktuelle und vergangene Sportengagement sowie die Einzel-Items zum aktuellen und vergangenen Leistungsniveau.

## *Ergebnisse*

### *Faktorielle Validität*

Die Analysen zur Struktur des MTI-D sind in Tabelle 1 dokumentiert. Alle Modelle wurden mittels MLR (Maximum Likelihood mit robusten Standardfehlerschätzungen) geschätzt. Als Rotationsverfahren wurde für Modell C eine schiefwinklige Geomin-Rotation eingesetzt und für Modell D eine orthogonale Bi-CF-Quartimax-Rotation (angelehnt an Marsh et al., 2009, bzw. Morin et al., 2016).

Tabelle 1 hier einfügen

Die Gütekriterien des Faktorenmodells erster Ordnung (Modell A) sprachen für eine akzeptable Passung (s. Tabelle 1). Auffällig waren die Faktorenkorrelationen: Sie betrogen

im Durchschnitt  $.60$  ( $.26 \leq r \leq .94$ ) und implizieren, dass die 12 Faktoren des MTI-D, ähnlich wie im Original (Middleton et al., 2005a), nur beschränkt eigenständige Aspekte erfassen.

Modell B, in dem diese Faktoreninterkorrelationen auf ein hierarchisch übergeordnetes Konstrukt zurückgeführt werden, schnitt im Vergleich mit Modell A deskriptiv wie inferenzstatistisch deutlich schlechter ab. Das weniger restriktive ESEM-Modell C mit 12 Faktoren erster Ordnung dagegen passte sehr gut zu den Daten – signifikant besser als Modell B. Für das BESEM-Modell D schließlich ergaben sich ebenfalls Werte (MLR- $\chi^2$  ( $240, N = 1\,122$ ) = 493.948,  $p < .001$ , CFI = .99, TLI = .97, SRMR = .01, RMSEA = .031), die auf eine sehr gute Datenpassung hinwiesen. Der Chi-Quadrat-Differenztest zeigte, dass Modell D signifikant besser passt als Modell C ( $TRd = 112.61, \Delta df = 24, p < .001$ ), doch im CFI, TLI und RMSEA unterschieden sich die Modelle kaum.

Angesichts dieser Befunde ließen sich sowohl das ESEM-Modell C als auch das BESEM-Modell D den weiteren Analysen zugrunde legen. Letztlich haben wir uns aus zwei Gründen für Modell D und damit für eine andere Faktorstruktur als jene entschieden, die Middleton et al. (2011) für das Original identifiziert haben: Primär entspricht es Middletons Konzeption des MTI; sekundär ist ein Bifaktormodell besonders gut geeignet, die mehrdimensionale Natur eines Instruments auszuloten und zu klären, inwiefern Mehrdimensionalität tatsächlich eine hinreichende Bedingung dafür ist, Subskalen zu bilden und sie einer Gesamtskala vorzuziehen (Reise et al., 2013).

Eine Durchsicht der standardisierten Faktorladungen von Modell D ließ erkennen (s. ESM 4), dass der Globalfaktor durch die 36 Items ( $.31 \leq |\lambda| \leq .79, M = .62$ ) gut definiert ist. Besonders hoch luden die Items 1, 13 und 25 sowie die Items 4, 16 und 28, die den Faktor *Zielbindung* (S1) bzw. *Selbstwirksamkeit* (S4) anvisieren. Relativ niedrig fielen die Ladungen der drei Items aus, die auf *Positive Vergleiche* (S8) ( $.31 \leq |\lambda| \leq .39$ ) zielen.

Die Ladungen der Items auf die spezifischen Faktoren waren – als direkte Folge der hohen Ladungen auf den Globalfaktor – relativ gering. Trotz Zulassung von Nebenladungen zeigte sich jedoch im Wesentlichen eine Einfachstruktur: Die Ladungen der Items auf die angezielten Faktoren waren meist klar höher als das Mittel und Maximum der Ladungen auf die nicht intendierten Faktoren. Problematisch erwiesen sich jedoch die Items 2, 7, 10, 13 und 25, die auf einen nicht intendierten Faktor höher luden als auf den intendierten. Auffällig ist Item 2, das auf den Faktor *Zielbindung* höher lud als die dafür vorgesehenen Items 1, 13 und 25.

Im Hinblick auf die Frage, wie gut die einzelnen spezifischen Faktoren durch die Items definiert sind, zeigte sich allgemein, dass innerhalb eines Faktors die Ladungen der intendierten Items im Durchschnitt höher ausfielen als das Mittel und Maximum aller übrigen Items. Wird (relativ willkürlich) ein Wert von .32 als Grenze einer hinreichenden Faktorladung gewählt (10 % gemeinsame Varianz nach Auspartialisierung des Globalfaktors), liegt die durchschnittliche Faktorladung der Items von zehn Faktoren darüber. Wir werten diesen Befund so, dass sie hinreichend gut durch die jeweils intendierten Items definiert sind. Auf die Faktoren *Zielbindung* (.23) und *Selbstwirksamkeit* (.30) trifft dies aber nicht zu.

#### *Reliabilität*

Weist die Strukturgleichungsmodellierung ein Instrument als mehrdimensional aus, wird dies oft so ausgelegt, dass die Skalen Facetten des Konstrukts erfassen, die in der Gesamtskala nicht repräsentiert sind, und dass die Gesamtskala kein einheitliches Konstrukt abbildet und darum nicht interpretierbar ist. Reise et al. (2013) dagegen erachten Mehrdimensionalität weder als notwendige noch hinreichende Bedingung dafür, Subskalen zu bilden und sie einer Gesamtskala vorzuziehen. Zur Entscheidungsfindung wurden auf der Basis von Modell D vier Formen von  $\omega$  berechnet (s. Tabelle 2 sowie ESM 4). Die innere Konsistenz ( $\omega$ ) betrug für die Gesamtskala .98, während sich die entsprechenden Werte ( $\omega_S$ ) für die Subskalen zwischen .81 und .91 bewegten. Weil für das MTI nur  $\alpha$ -Koeffizienten vorliegen, wurde auch dieses Maß geschätzt (Tabelle 2). Die Werte für  $\omega$ ,  $\omega_S$  und  $\alpha$  wichen dabei kaum voneinander ab (s. ESM 3 für die Befunde zur Retest-Reliabilität des MTI-D).

Der Globalfaktor *Mental Toughness* wies mit  $\omega_H = .94$  eine hohe Reliabilität auf, während die entsprechenden Werte für die spezifischen Faktoren im Durchschnitt .35 ( $.07 \leq \omega_{HS} \leq .77$ ) betrugten und deutlich unter der internen Konsistenz des Globalfaktors lagen. Relativ hohe Werte waren für die spezifischen Faktoren *Potential*, *Positive Vergleiche* und *Aufgabenvalenz* zu beobachten ( $\omega_{HS} \geq .53$ ). Anhand des Quotienten  $\omega_{HS}/\omega_S$  lässt sich zudem der Varianzanteil eines spezifischen Faktors schätzen, der vom Globalfaktor unabhängig ist (Rodriguez et al., 2016; Tabelle 2). Dieser Quotient betrug im Mittel .41 und bewegte sich zwischen .08 (*Zielbindung*) und .84 (*Positive Vergleiche*).

Tabelle 2 hier einfügen

### *Messinvarianz*

Die Befunde zur Messinvarianz sind in Tabelle 1 zusammengefasst. Die Schätzung von Modell D für die Gruppe der Vereinssportlerinnen und Vereinssportler bzw. Sportstudierenden ergab Werte, die auf eine gute (D1) bzw. eine akzeptable Passung hinweisen (D2). Damit wird konfigurale Messinvarianz angenommen. Für das Basismodell (D3) zur Prüfung der metrischen Messinvarianz wurden in beiden Gruppen alle Parameter frei geschätzt. Das Ausmaß der Veränderung von CFI bzw. RMSEA durch die Gleichsetzung der Faktorenladungen in den Gruppen (Basismodell D3 vs. metrisches Modell D4) und in einem weiteren Schritt durch die Gleichsetzung der sich entsprechenden Item-Mittelwerte in den Gruppen (metrisches Modell D4 vs. skalares Modell D5) blieb jeweils unterhalb der zurate gezogenen Grenzwerte. Damit können auch die metrische und skalare Messinvarianz als gegeben betrachtet werden. Die entsprechenden Analysen wurden auch für zwei Geschlechts- und Altersgruppen durchgeführt. Auch für diese beiden Merkmale kann konfigurale, metrische und skalare Messinvarianz angenommen werden.

### *Konstrukt- und Kriteriumsvalidität*

Mit Blick auf die Konstruktvalidität sind eine allgemeine Hypothese und vier spezifische Hypothesen formuliert worden. Was die Allgemeine betrifft, zeigt Tabelle 2, dass der Globalfaktor *Mentale Stärke* mit sechs TOPS-D-Skalen die erwarteten Korrelationen aufwies ( $.30 \leq |r| \leq .46$ ). Mit den restlichen drei Skalen (*Selbstgespräch*, *Zielfokussierung* und *Visualisierung*) korrelierte er jedoch nicht signifikant. Was die spezifischen Hypothesen und damit einzelne spezifische Faktoren angeht, aus denen der Globalfaktor auspartialisiert worden ist, ergaben sich weniger ausgeprägte Korrelationen ( $.16 \leq |r| \leq .31$ ), die sich mehrheitlich als signifikant erwiesen ( $\alpha = .05$ ): *Stressminimierung* und *Positive Einstellung* korrelierten mit *Emotionale Kontrolle* und *Negative Kognitionen*, und *Aufgabenbezogene Aufmerksamkeit* hing wie erwartet mit *Aufmerksamkeitskontrolle* zusammen. Die Korrelationen zwischen den beiden spezifischen MTI-D-Faktoren *Selbstwirksamkeit* und *Selbstkonzept mentaler Stärke* einerseits und *Negative Kognitionen* andererseits erreichten die gesetzte Signifikanzschwelle jedoch nicht (-.08 bzw. -.15). Es ergaben sich noch weitere Zusammenhänge zwischen den spezifischen MTI-D-Faktoren und einzelnen TOPS-D-Skalen ( $|r_{\max}| = .38$ ), zu denen aber keine Hypothesen formuliert wurden.

Die erwarteten Korrelationen zeigten sich auch zwischen dem Globalfaktor *Mental Toughness* und dem Sportengagement bzw. Leistungsniveau. Die implizite Vermutung, dass

die Beziehungen zu gegenwartsnahen Kriterien höher ausfallen als die Beziehungen zu weiter zurückliegenden, fand in den Daten keine Entsprechung. Auf der Ebene der spezifischen MTI-D-Faktoren ergaben sich erneut einige signifikante Korrelationen. Vor allem die Faktoren *Zielbindung* und *Positive Vergleiche* korrelierten mit den vier Kriterien. Dazu wurden jedoch keine Hypothesen formuliert, und auch post hoc ist eine Interpretation dieser Befunde nicht offensichtlich.

### *Diskussion und Ausblick*

Das Ziel dieser Studie war es, die Testgüte des MTI-D bezüglich faktorieller Struktur, Messinvarianz, Reliabilität sowie Konstrukt- und Kriteriumsvalidität zu analysieren.

In struktureller Hinsicht offenbarte ein konfirmatorisches Modell mit 12 Faktoren erster Ordnung (CFA, Modell A) eine akzeptable, aber keine so gute Datenpassung, wie Middleton et al. (2005a) dies für das Original berichteten. Zudem korrelierten die Faktoren untereinander stark, was mit Blick auf ihre Differenzierungsfähigkeit ungünstig ist. Als nächstes wurde ein konfirmatorisches Modell zweiter Ordnung (Modell B) geschätzt – und verworfen, das diese Korrelationen auf einen übergeordneten Faktor zurückführt. Ob es sich hier um ein Spezifikum des MTI-D handelt, ist ungewiss, da entsprechende Befunde zum Original fehlen. Ein exploratorisches Faktorenmodell erster Ordnung (ESEM, Modell C) sowie ein exploratorisches Bifaktormodell (BESEM, Modell D) wiesen jeweils einen sehr guten Fit auf. Dieses BESEM-Modell (mit orthogonalen Faktoren) passte dabei nicht nur etwas besser als das ESEM-Modell (mit korrelierten Faktoren), sondern war auch dem CFA-Modell erster Ordnung (Modell A) deutlich überlegen, das sich für das Original aufgedrängt hat.

Befunde, die von denjenigen zum Original-Instrument abweichen, kommen oft ungelegen – in diesem Fall aber nicht, denn das Bifaktormodell entspricht der theoretischen Vorstellung, auf dem der MTI gründet (Middleton et al., 2011), die aber empirisch nicht geprüft worden ist. Unsere Befunde, die unter Anwendung des relativ neuen ESEM-Ansatzes und des wiederentdeckten Bifaktormodells gewonnen wurden, deuten nun darauf hin, dass dem MTI-D ein globaler Faktor sowie 12 davon unabhängige spezifische Faktoren zugrunde liegen.<sup>2</sup>

---

<sup>2</sup> Wer mit A-priori-Skalen arbeitet, mag sich für deren Zusammenhang mit den Faktorwerten (Modell D) interessieren (s. ESM 4).

Das BESEM-Modell überzeugte zwar bezüglich der Gesamtstruktur, doch die Faktorladungsmatrix ließ auch fünf suboptimale Items erkennen: Item 2 („Es ist für mich wichtig zu versuchen, das Beste zu geben, was ich kann“, *SNB*), Item 7 („Wenn Schwierigkeiten auftreten, arbeite ich weiter und akzeptiere keinen Misserfolg“, *DV*), Item 10 („Aufgrund meiner Erfahrung in diesem Bereich erbringe ich auf meinem jetzigen Niveau gute Leistungen“, *VA*), Item 13 („Mein Einsatz für meine Ziele ist stark“, *ZB*) und Item 25 („Was auch immer passiert, ich setze mich für meine Ziele ein“, *ZB*). Auch Middleton et al. (2005a, 2011) fanden im Original für die Items 2, 7 und 10 etwas geringere Ladungen, doch waren in unserer Studie die Ladungen dieser drei Items auf den jeweils angezielten Faktor mit .09, .08 und .08 (unter Kontrolle des Globalfaktors) deutlich unbefriedigend. Hier könnte die Weiterentwicklung des MTI-D ansetzen: Im Falle der Items 2, 10 und 13 könnte eine sprachliche Glättung die erhoffte Wirkung bringen. Im Falle von Item 2, das wie das im Fragebogen vorangehende Item 1 stark, aber nicht intendiert auf den Faktor *Zielbindung* ( $\lambda = .68$ ) lud, lässt sich zusätzlich ein Positionseffekt vermuten, der allenfalls durch Eisbrecher-Items (Lienert & Raatz, 1998) zu entschärfen ist.

Die innere Konsistenz der Gesamtskala und der 12 Subskalen ( $\omega = .98$ ;  $.81 \leq \omega_S \leq .91$ ) bewegte sich in einem ähnlich hohen Bereich wie jene der englischsprachigen MTI-Skalen ( $.83 \leq \alpha \leq .93$ ; Middleton et al., 2005a, 2011) und übertraf die Konventionen für eine Anwendung in der Gruppen- und Individualdiagnostik. Die Analysen zur Retest-Reliabilität der MTI-D-Skalen über zwei, fünf und sieben Monate (s. ESM 3) ergaben eine Korrelation der Gesamtskala für die drei Intervalle von .70, .74 und .78; die entsprechenden gemittelten Korrelationen der Subskalen beliefen sich auf .66, .68 und .68. Ohne vergleichbare Analysen zum Original, sind die Befunde für das MTI-D nicht gut einzuordnen. Insgesamt deuten sie aber auf relativ messgenaue Skalen oder ein relativ stabiles Konstrukt hin.

Reise et al. (2013) folgend kann omega hierarchical als weiterer Test für die Angemessenheit einer Faktorenlösung aufgefasst werden. Für den Globalfaktor ( $\omega_H$ ) resultierte ein hoher Wert von .94, während sich für die spezifischen Faktoren ( $\omega_{HS}$ ) – absolut und relativ zu  $\omega_S$  – geringe Werte ergaben ( $M = .35$ ). Wie das Beispiel des Faktors *Zielbindung* zeigt (.07), sind geringe Werte für  $\omega_{HS}$  komplexen Faktorladungsmatrizen geschuldet: So luden die Items 1, 13 und 25 mit Werten von .68, .79 und .76 zum einen hoch auf den Globalfaktor. Zum anderen wiesen diese drei Items unter Kontrolle des Globalfaktors relativ hohe Ladungen auf einen nicht intendierten Faktor auf (-.16, -.17 und -.21), die zum Teil



höher ausfielen als die intendierten Ladungen auf den Faktor *Zielbindung* (.44, .14 und .08). Übers Ganze gesehen genügen *Zielbindung* und sechs weitere Skalen den Ansprüchen von Reise et al. (2013,  $\omega_{HS} = .50$ ) nicht, nämlich *Streben nach Bestleistung*, *Selbstwirksamkeit*, *Aufgabenbezogene Aufmerksamkeit*, *Durchhaltevermögen*, *Vertrautheit mit der Aufgabe* und *Selbstkonzept mentaler Stärke*. Genügend sind dagegen *Potential*, *Positive Vergleiche* und *Aufgabenvaleanz* und bei großzügiger Auslegung auch *Stressminimierung* und *Positive Einstellung*.

Ob daraus zu schließen ist, dass diese fünf oder sieben spezifischen Faktoren in ihrer jetzigen Form entbehrlich sind, ist jedoch unklar, denn die Diskussion um Richtlinien zur Beurteilung von  $\omega_{HS}$  hat noch kaum begonnen. So verwenden Arens und Morin (2016) unbesehen die gängigen Konventionen für  $\alpha$  und  $\omega$ , um die korrigierte Reliabilität von spezifischen Faktoren ( $\omega_{HS}$ ) zu beurteilen. Und die Gruppe um Reise greift in einer aktuellen Arbeit (Rodriguez et al., 2016) ihre Empfehlungen nicht mehr auf, sondern lassen empirische Befunde sprechen – gewonnen an 50 (der noch wenigen) Studien, in denen Instrumente mit dem Bifaktormodell untersucht worden sind: Über 151 Subskalen hinweg betrachtet lag der Mittelwert von  $\omega_{HS}$  bei .27 ( $SD = .17$ ), und nur 20 (13 %) Skalen erreichten das vorgeschlagene Minimum. Aus forschungspraktischer Sicht bleibt anzufügen, dass im Kontext der Strukturgleichungsmodellierung Werte auf messfehlerbereinigten latenten Variablen als perfekt reliabel betrachtet werden können (Arens & Morin, 2016), auch wenn die auf der Basis der standardisierten Faktorladungen berechneten  $\omega_H$  oder  $\omega_{HS}$  gering sind.

Die Messinvarianz des MTI-D wurde in drei Hinsichten geprüft. Die Gruppen der Vereinssportlerinnen bzw. -sportler und Sportstudierenden sowie die zwei Geschlechts- und Altersgruppen waren in puncto Faktorstruktur, -ladungen und Itemmittelwerte parallel. Wenngleich weitere Abklärungen zur Messinvarianz wünschenswert sind (z. B. Leistungsniveau), deutet sich an, dass das MTI-D in verschiedenen Subgruppen äquivalent misst und entsprechend angewendet werden kann.

Die Analysen zur Konstrukt- und Kriteriumsvalidität des MTI-D ergaben, dass der Globalfaktor mit sechs TOPS-D-Skalen zusammenhing – nicht aber mit den Skalen *Selbstgespräch*, *Zielfokussierung* und *Visualisierung*. Dieses Befundmuster ist allenfalls damit erklärbar, dass diese drei TOPS-D-Skalen vorrangig Techniken erfassen, die Athletinnen und Athleten anwenden, um in den optimalen Leistungszustand zu gelangen. Die erwartungskonform korrelierenden Skalen *Aktivierung*, *Emotionale Kontrolle*,

*Aufmerksamkeitskontrolle*, *Negative Kognitionen* und *Automatisierung* dagegen beziehen sich jeweils auf eine Fertigkeit und erfassen das Ausmaß, in dem es Athletinnen und Athleten unabhängig vom Einsatz einer bestimmten Technik gelingt, sich in den optimalen Leistungszustand zu versetzen. Die Nullkorrelationen würden dann bedeuten, dass mentale Stärke nicht an die bewusste und erfolgreiche Anwendung dieser Techniken gebunden ist. Allerdings passt zu dieser Interpretation nicht recht, dass *Mental Toughness* und *Entspannung* korrelieren, die vierte TOPS-D-Dimension, die eine mentale Technik erfasst.

Zwischen dem Globalfaktor des MTI-D und dem Sportengagement bzw. Leistungsniveau waren die erwarteten Korrelationen zu beobachten. Die Unterscheidung des zeitlichen Bezugsrahmens (1 Jahr vs. 10 Jahre) ergab kein klares Bild, denn die Korrelationen mit den gegenwartsnahen Kriterien wichen kaum von jenen mit den zurückliegenden ab. Da die verwendeten Maße bezüglich Gegenwartsnähe und Länge des Zeitintervalls nicht trennscharf sind, haben wir dazu jedoch keine explizite Erwartung geäußert, etwa im Sinne eines Primär- oder Rezenzeffekts.

Auch auf der Ebene der spezifischen MTI-D-Faktoren fanden die meisten Hypothesen in den Daten eine Entsprechung. Beispielsweise korrelierten die MTI-D-Faktoren *Stressminimierung* und *Positive Einstellung* mit der TOPS-D-Skala *Negative Kognitionen*. Diese Zusammenhänge sind bemerkenswert, da sie auf einer kleinen Stichprobe gründen und um den Effekt des Globalfaktors bereinigt sind. Zwischen den spezifischen MTI-D-Faktoren einerseits und dem TOPS-D, Sportengagement und Leistungsniveau andererseits ergaben sich aber auch überraschende, nämlich negative Zusammenhänge. Dies betrifft vorab die MTI-D-Faktoren *Durchhaltevermögen* und *Zielbindung*. Weil in diesen spezifischen Faktoren wenig reliable Varianz steckt, mögen diese Befunde zufallsbedingt sein. Es könnte sich darin aber auch andeuten, dass diese Faktoren nach Auspartialisierung des Globalfaktors nur noch einen maladaptiven Anteil von Durchhaltevermögen und Zielbindung beinhalten: ein unzweckmäßiges Nicht-locker-Lassen oder situationsinadäquates Beharren auf Zielen. Insgesamt aber haben diese Analysen vorläufige Hinweise auf die Konstrukt- und Kriteriumsvalidität des MTI-D ergeben.

Wenn ein Fazit aus den testpsychologischen Analysen eines Verfahrens gezogen wird, kommt denjenigen zur faktoriellen Validität eine herausragende Bedeutung zu (Gucciardi et al., 2012). Es ist darum festzuhalten, dass sich klare Indizien für die Mehrdimensionalität des MTI-D ergeben haben. In Erweiterung der Arbeiten zum englischsprachigen Original

(zusammenfassend Middleton et al., 2011) legen die Befunde den Schluss nahe, dass dem MTI-D ein gemeinsamer Faktor *Mental Toughness* und mehrere spezifische Faktoren zugrunde liegen, die Aspekte von mentaler Stärke abdecken, die im Globalfaktor nicht enthalten sind. Die Befunde machen aber auch darauf aufmerksam, dass noch unklar ist, ob die diagnostische Nutzung des MTI-D in der aktuellen Form – neben dem globalen Faktor – tatsächlich auf allen spezifischen Faktoren gründen kann.

Wird an den 12 Facetten festgehalten, muss eine Weiterentwicklung des MTI-D an jenen Faktoren ansetzen, die noch wenig spezifische Information erfassen, namentlich *Zielbindung, Streben nach Bestleistung, Selbstwert, Aufgabenbezogene Aufmerksamkeit, Durchhaltevermögen, Vertrautheit mit der Aufgabe* und *Selbstkonzept mentaler Stärke*. Es wird sich weisen, ob sich Items generieren lassen, mit denen das Spezifische dieser Faktoren besser erfasst werden kann. Sollte dies nicht realisierbar sein, ist im anderen Extrem die Auswertung des MTI-D auf den Generalfaktor zu beschränken. Doch ein 36-Item-Instrument ist wenig ökonomisch, und es stellt sich die Frage nach einer optimalen Kurzform.

Auf der Basis einer weiterentwickelten Version des MTI-D sind generell Studien mit größeren Stichproben erstrebenswert und speziell solche mit größeren Gruppen von Athletinnen und Athleten aus Sportarten mit ähnlichen psychophysischen Anforderungen. Zum einen lässt sich dann auch die Konstrukt- und Kriteriumsvalidität mittels Strukturgleichungsmodellierung analysieren. Zum andern könnten Athletinnen und Athleten z. B. aus ausdauerorientierten oder ästhetisch-kompositorischen Sportarten miteinander verglichen werden, um Hypothesen über die differenzielle Leistungsrelevanz von MTI-D-Faktoren wie *Stressminimierung* oder *Durchhaltevermögen* zu prüfen. Solche Studien wären wichtig, um zu klären, ob die differenzierte Erfassung von mentaler Stärke diagnostisch tatsächlich gewinnbringend ist. Sinnvoll sind auch Stichproben aus leistungsbezogenen, aber sportfernen Kontexten (z. B. Arbeit, Rehabilitation), zumal das Instrument sportunspezifisch eingesetzt werden kann. Zur weiteren Validierung des MTI-D ist es schließlich wünschenswert, psychologische Konstrukte wie Selbstwirksamkeit, Optimismus oder Resilienz sowie ein Instrument wie den MTQ48 heranzuziehen, der explizit mentale Stärke erfassen will.

Ein zentrales Anwendungsfeld des MTI-D ist der leistungsorientierte Sport, weil mentale Stärke als wichtiger Faktor für das Erbringen guter Leistungen gilt. Dabei lässt sich der MTI-D in der Forschung einsetzen wie auch in der individuellen Beratung.

Beispielsweise kann er vor dem Hintergrund des Choking-under-Pressure-Phänomens (Beilock & Gray, 2007) oder der Verletzungsrehabilitation helfen, diagnostische Informationen zu sammeln, auf deren Basis sich adressatengerechte Interventionen gestalten lassen, mit denen die psychischen Voraussetzungen für die Erbringungen sportlicher Leistungen oder die Bewältigung von Laufbahnzäsuren optimiert werden. Wird ein differenziertes Bild angestrebt und die 12 Subskalen des MTI-D eingesetzt, ist zu bedenken, dass die Skalen nicht völlig eigenständige Aspekte von mentaler Stärke erfassen. Geht es um eine weniger differenzierte Diagnose, ist grundsätzlich auch die Gesamtskala einsetzbar, allerdings ist sie relativ lang.

Damit ein Erhebungsinstrument den zentralen Testgütekriterien genügt, sind meist mehrere Entwicklungsschritte nötig. In Bezug auf das MTI-D ist hervorzuheben, dass es ähnlich gut oder schlecht misst wie das Original. Die Punkte, die wir im MTI-D als problematisch ansehen, dürften kaum in unserer Übersetzung oder Stichprobe begründet sein. Sie sind eher unseren weiterführenden Analysen geschuldet, die Schwächen sichtbar machen, die sich auch im Original abzeichnen, dort aber nicht aufgegriffen wurden (z. B. Faktorkorrelationen). Ferner ist zu berücksichtigen, dass die mittlerweile in deutscher Sprache verfügbare Version des MTQ48 (Gerber et al., 2013) zwar eine Alternative darstellt, aber noch einer gründlichen Analyse harret und kaum am Ende ihrer Entwicklung angelangt ist.

## **Elektronische Supplemente (ESM)**

Die elektronischen Supplemente sind mit der Online-Version dieses Artikels verfügbar über [vom Verlag ergänzen]

*ESM 1.* Tabelle 1 (ESM 1 Beispielitems.pdf).

Die Tabelle gibt Beispielitems des MTI-D wieder.

*ESM 2.* Abbildung 1 (ESM 2 MTI-Modelle.pdf).

Graphische Darstellung unterschiedlicher MTI-D-Modelle.

*ESM 3.* Retest-Reliabilität (ESM 3 Retest-Reliabilität.pdf).

Methodisches Vorgehen und Ergebnisse zur Retest-Reliabilität des MTI-D.

*ESM 4.* Tabelle 3 (ESM 4 Faktorladungen.pdf).

Die Tabelle enthält standardisierte Faktorladungen, Residualvarianzen und Reliabilitätskoeffizienten (McDonalds  $\omega$ ) des exploratorischen Bifaktormodells (BESEM, Modell D).

*ESM 5.* Tabelle 4 (ESM 5 Korrelationen.pdf).

Die Tabelle enthält die Korrelationen zwischen Faktorwerten und Skalen des MTI-D.

## Literatur

- Arens, A. K. & Morin, A. J. S. (2016). Improved representation of the Self-Perception Profile for Children through bifactor exploratory structural equation modeling. *American Educational Research Journal*. Advance online publication.  
doi:10.3102/0002831216666490
- Asparouhov, T. & Muthén, B. (2009). Exploratory structural equation modeling. *Structural Equation Modeling*, 16, 397–438.
- Bandura, A. (1997). *Self-efficacy: The exercise of control*. New York, NY: Freeman.
- Beilock, S. L. & Gray, R. (2007). Why do athletes choke under pressure? In G. Tenenbaum & R. C. Eklund (Eds.), *Handbook of sport psychology* (3rd ed., pp. 425–444). Hoboken, NJ: Wiley.
- Bortz, J. & Schuster, C. (2010). *Statistik für Human- und Sozialwissenschaftler* (7. Aufl.). Berlin: Springer.
- Brown, T. A. (2015). *Confirmatory factor analysis*. London: Guilford.
- Chen, F. F. (2007). Sensitivity of goodness of fit indexes to lack of measurement invariance. *Structural Equation Modeling*, 14, 464–504.
- Clough, P., Earle, K. & Sewell, D. (2002). Mental toughness: The concept and its measurement. In I. Cockerill (Ed.), *Solutions in sport psychology* (pp. 32–46). London: Thomson Learning.
- Connaughton, D. & Hanton, S. (2010). (2009). Mental toughness in sport: Conceptual and practical issues. In S. D. Mellalieu & S. Hanton (Eds.), *Advances in applied sport psychology: A review* (pp. 317–346). New York, NY: Routledge.
- Gerber, M. (2011). Mentale Toughness im Sport. *Sportwissenschaft*, 41, 283–299.
- Gerber, M., Kalak, N., Lemola, S., Clough, P. J., Perry, J. L., Pühse, U. et al. (2013). Are adolescents with high mental toughness levels more resilient against stress? *Stress and Health*, 29, 164–171.
- Gucciardi, D. F., Hanton, S. & Mallett, C. J. (2012). Progressing measurement in mental toughness: A case example of the Mental Toughness Questionnaire 48. *Sport, Exercise, and Performance Psychology*, 1, 194–214.
- Gucciardi, D. F., Mallett, C. J., Hanrahan, S. J. & Gordon, S. (2011). Measuring mental toughness. Current status and future directions. In D. F. Gucciardi & S. Gordon (Eds.), *Mental toughness in sport: Developments in theory and research* (pp. 108–132). New

York, NY: Routledge.

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

Lienert, G. A. & Raatz, U. (1998). *Testaufbau und Testanalyse* (6. Aufl.). Weinheim: Psychologie Verlags Union.

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

Marsh, H. W., Muthén, B., Asparouhov, T., Lüdtke, O., Robitzsch, A., Morin, A. J. S. & Trautwein, U. (2009). Exploratory structural equation modeling, integrating CFA and EFA: Application to students' evaluations of university teaching. *Structural Equation Modeling*, 16, 439–476.

McDonald, R. P. (1999). *Test theory: A unified treatment*. Mahwah, NJ: Erlbaum.

Middleton, S. C., Marsh, H. W., Martin, A. J., Richards, G. E. & Perry, C. (2004a, July). *Developing the Mental Toughness Inventory (MTI)*. Paper presented at the 3rd International Biennial SELF Research Conference on Self-Concept, Motivation and Identity: Where to go from here? Berlin. Zugriff unter <http://researchdirect.uws.edu.au/islandora/object/uws%3A7274>

Middleton, S. C., Marsh, H. W., Martin, A. J., Richards, G. E. & Perry, C. (2004b, July). *Discovering mental toughness: A qualitative study of mental toughness in elite athletes*. Paper presented at the 3rd International Biennial SELF Research Conference on Self-Concept, Motivation and Identity: Where to go from here? Berlin. Zugriff unter [http://www.sectiononewrestling.com/discovering\\_mental\\_toughness.pdf](http://www.sectiononewrestling.com/discovering_mental_toughness.pdf)

Middleton, S. C., Marsh, H. W., Martin, A. J., Richards, G. E. & Perry, C. (2005a). *Developing a test for mental toughness: The Mental Toughness Inventory (MTI)*. Paper presented at the Joint AARE/NZARE Conference, Parramatta, Australia. Zugriff unter <http://www.aare.edu.au/data/publications/2005/mid05310.pdf>

Middleton, S. C., Marsh, H. W., Martin, A. J., Richards, G. E. & Perry, C. (2005b). *Making the leap from good to great: Comparisons between sub-elite and elite athletes on mental toughness*. Paper presented at the Joint AARE/NZARE Conference, Parramatta, Australia. Zugriff unter <http://www.aare.edu.au/data/publications/2005/mid05301.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.
- Middleton, S. C., Martin, A. J. & Marsh, H. W. (2011). Development and validation of the Mental Toughness Inventory (MTI): Construct validation approach. In D. F. Gucciardi & S. Gordon (Eds.), *Mental toughness in sport: Developments in theory and research* (pp. 91–107). New York, NY: Routledge.
- Morin, A. J. S., Arens, A. K. & Marsh, H. W. (2016). A bifactor exploratory structural equation modeling framework for the identification of distinct sources of construct-relevant psychometric multidimensionality. *Structural Equation Modeling*, *23*, 116–139.
- Muthén, L. K. & Muthén, B. O. (1998–2015). *Mplus user's guide* (7th ed.). Los Angeles, CA: Muthén & Muthén.
- Reise, S. P. (2012). The rediscovery of bifactor measurement models. *Multivariate Behavioral Research*, *47*, 667–696.
- Reise, S. P., Bonifay, W. E. & Haviland, M. G. (2013). Scoring and modeling psychological measures in the presence of multidimensionality. *Journal of Personality Assessment*, *95*, 129–140.
- Richards, H. (2011). Coping and mental toughness. In D. Collins, A. Abbott & H. Richards (Eds.), *Performance psychology: A practitioner's guide* (pp. 281–300). Edinburgh: Elsevier.
- Rodriguez, A., Reise, S. P. & Haviland, M. G. (2016). Evaluating bifactor models: Calculating and interpreting statistical indices. *Psychological Methods*, *21*, 137–150.
- Satorra, A. & Bentler, P. M. (2001). A scaled difference chi-square test statistic for moment structure analysis. *Psychometrika*, *66*, 507–514.
- Schmid, J., Birrer, D., Kaiser, U. & Seiler, R. (2010). Psychometrische Eigenschaften einer deutschsprachigen Adaptation des Test of Performance Strategies (TOPS): Ein Instrument zur Erfassung psychischer Fertigkeiten im Sport. *Zeitschrift für Sportpsychologie*, *17*, 50–62.
- Schmid, J. & Gubelmann, H. (1999). A psychometric evaluation of a German version of the Loehr Test and the Sports Emotional-Reaction Profile. In V. Hosek, P. Tilinger & L. Bilek (Eds.), *Psychology of sport and exercise: Enhancing the quality of life. Proceedings of the 10th European Congress of Sport Psychology – FEPSAC* (Part 2, pp. 148–150). Prague, Czech Republic: Charles University.



- Shavelson, R. J., Hubner, J. J. & Stanton, G. C. (1976). Self-concept: Validation of construct interpretations. *Review of Educational Research*, 48, 407–441.
- Sheard, M. (2012). *Mental Toughness: The mindset behind sporting achievement* (2nd ed.). London: Routledge.
- 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, 186–193.
- Vallerand, R. J. (1989). Vers une méthodologie de validation trans-culturelle de questionnaires psychologiques: Implications pour la recherche en langue française. *Canadian Psychology/Psychologie Canadienne*, 30, 662–680.

Danksagung

Dieses Forschungsprojekt wurde durch die Eidgenössische Sportkommission (ESK)  
finanziell unterstützt.

Tabelle 1. Prüfung der Faktorstruktur (Gesamtstichprobe,  $N = 1\,122$ ) und der Messinvarianz des MTI-D für Sport-, Geschlechts- und Altersgruppen

Modell	MLR- $\chi^2$	df	CFI	TLI	SRMR	RMSEA	90 % KI RMSEA	CFit
A CFA-Modell erster Ordnung	2 029.945***	528	.929	.915	.046	.050	[.048, .053]	.397
B HCFA-Modell zweiter Ordnung	3 405.825***	582	.867	.856	.084	.066	[.064, .068]	< .001
C <sup>a</sup> ESEM-Modell	614.003***	264	.983	.961	.010	.034	[.031, .038]	> .999
D <sup>a</sup> BESEM-Modell	493.948***	240	.988	.969	.009	.031	[.027, .035]	> .999
Messinvarianz <sup>b, c</sup>								
D1 Vereinssportler ( $n = 914$ )	470.633***	240	.987	.966	.009	.032	[.028, .037]	> .999
D2 Sportstudierende ( $n = 208$ )	398.249***	240	.964	.905	.016	.056	[.046, .066]	.142
D3 Basismodell	875.590***	480	.983	.955	.010	.038	[.034, .042]	> .999
D4 Metrisches Modell	1 363.748***	779	.975	.959	.023	.037	[.033, .040]	> .999
D5 Skalares Modell	1 393.491***	802	.974	.960	.032	.036	[.033, .039]	> .999
Frauen ( $n = 465$ )								
Männer ( $n = 657$ )	500.314***	240	.981	.950	.009	.041	[.036, .046]	.999
Basismodell	821.186***	480	.985	.960	.010	.036	[.031, .040]	> .999
Metrisches Modell	1 240.318***	779	.979	.966	.023	.032	[.029, .036]	> .999
Skalares Modell	1 274.022***	802	.979	.966	.024	.032	[.029, .036]	> .999
Jüngere Altersgruppe <sup>d</sup> (< 22 Jahre; $n = 547$ )								
Ältere Altersgruppe ( $\geq 22$ Jahre; $n = 575$ )	499.142***	240	.977	.939	.011	.043	[.038, .049]	.980
Basismodell	913.360***	480	.981	.949	.010	.040	[.036, .044]	> .999
Metrisches Modell	1 207.862***	779	.981	.969	.020	.031	[.028, .035]	> .999
Skalares Modell	1 265.661***	802	.980	.968	.021	.032	[.028, .035]	> .999

Anmerkungen: MLR = Robuste Maximum-Likelihood-Schätzung nach Yuan-Bentler; CFI = Comparative Fit Index; TLI = Tucker-Lewis Index; SRMR = Standardized Root Mean Square Residual; RMSEA = Root Mean Square Error of Approximation; 90 % KI = 90%-Konfidenzintervall für RMSEA; CFit =  $p$ -Wert des Close-Fit-Tests. (H)CFA = (hierarchical) confirmatory factor analysis; (B)ESEM = (bifactor) exploratory structural equation modeling.

<sup>a</sup> Wegen Schätzproblemen wurde die (geringfügig) negative Residualvarianz des Items 34 auf einen Wert > 0 gesetzt. <sup>b</sup> Allen Abklärungen zur Messinvarianz liegt Modell D zugrunde. <sup>c</sup> Infolge von Schätzproblemen bei den Abklärungen zur Messinvarianz bezüglich des Merkmals Vereinssport vs. Sportstudium wurde die (geringfügig) negative Residualvarianz der Items 1, 24, 31, 34 oder 36 auf einen Wert > 0 restringiert. <sup>d</sup> Die Altersgruppen wurden via Mediansplit gebildet.

\*\*\* $p < .001$ .

*Tabelle 2. Reliabilität des MTI-D und Rangkorrelationen zwischen MTI-D und TOPS-D, Sportengagement sowie Leistungsniveau*

MTI-D-Faktorwerte (BESEM- Modell D) bzw. -Skalen	Interne Konsistenz ( <i>N</i> = 1 122)				TOPS-D-Skalen ( <i>n</i> = 93)									Sport- engagement ( <i>n</i> = 115)		Leistungs- niveau ( <i>n</i> = 115)	
	$\alpha$	$\omega_s$	$\omega_{HS}$	$\omega_{HS}/\omega_s$	SG	EK	AUT	ZF	VIS	AKT	ENT	AUF	NK	1 J. <sup>d</sup>	10 J. <sup>e</sup>	1 J.	10 J.
Globalfaktor/Gesamtskala	.96	.98 <sup>a</sup>	.94 <sup>b</sup>	.96 <sup>c</sup>	.07	.54	.28	.08	.05	.47	.42	.46	-.45	.40	.40	.31	.27
Zielbindung	.86	.87	.07	.08	.03	-.34	-.18	-.05	.01	-.16	-.34	-.28	.21	-.23	-.24	-.15	-.20
Streben nach Bestleistung	.82	.86	.20	.23	-.06	-.19	-.07	-.01	.05	.02	-.04	-.17	.17	.01	-.01	-.02	-.02
Stressminimierung	.80	.81	.37	.45	.06	.27	.24	-.01	.06	.08	.13	.37	-.20	-.05	-.15	-.11	-.08
Selbstwirksamkeit	.87	.86	.13	.15	-.04	-.01	.06	.00	-.04	.20	.16	.06	-.08	.13	.02	.07	.09
Aufgabenbezogene Aufmerksamkeit	.88	.89	.31	.35	-.07	.22	.01	-.03	-.10	-.13	.11	.23	-.09	.02	.05	.00	.04
Potential	.89	.89	.53	.60	.19	.02	.10	-.08	.18	.07	.05	.01	-.04	.01	.10	.20	.04
Durchhaltevermögen	.79	.82	.19	.24	-.01	-.23	-.19	.23	.10	.01	-.20	-.14	.13	-.18	-.14	-.12	-.04
Positive Vergleiche	.91	.91	.77	.84	.07	-.09	.04	-.01	-.13	-.11	.07	-.06	.06	.20	.17	.23	.18
Aufgabenvaleanz	.90	.91	.53	.58	.06	-.04	.01	.16	-.09	.08	-.14	.06	.22	.08	.13	.04	-.07
Vertrautheit mit der Aufgabe	.77	.83	.23	.28	.20	-.09	-.03	.15	.12	-.01	-.19	.00	.11	.15	.09	.05	.09
Positive Einstellung	.86	.87	.35	.40	.04	.23	.01	.03	.13	.22	.22	.17	-.34	-.14	-.04	-.07	-.19
Selbstkonzept mentaler Stärke	.88	.89	.28	.31	-.03	.10	.10	-.09	.06	.20	.38	.02	-.15	.08	.09	.15	.09

*Anmerkungen.* MTI-D: Mental Toughness Inventory (deutschsprachige Version). TOPS-D: Test of Performance Strategies (deutschsprachige Version). BESEM: Bifactor exploratory structural equation modeling.  $\alpha$  = Cronbachs  $\alpha$ ;  $\omega$  = omega;  $\omega_s$  = omega subscale;  $\omega_H$  = omega hierarchical;  $\omega_{HS}$  = omega hierarchical subscale (Rodriguez et al., 2016). TOPS-D: SG: Selbstgespräch ( $\alpha$  = .90), EK: Emotionale Kontrolle (.82), AUT: Automatisierung (.84), ZF: Zielfokussierung (.79), VIS: Visualisierung (.80), AKT: Aktivierung (.68), ENT: Entspannung (.79), AUF: Aufmerksamkeitskontrolle (.69), NK: Negative Kognitionen (.74).

<sup>a</sup>  $\omega$ . <sup>b</sup>  $\omega_H$ . <sup>c</sup>  $\omega_H/\omega$ . <sup>d</sup> Bezogen auf den 1-Jahres-Zeitraum vor dem Erhebungszeitpunkt («aktuell»). <sup>e</sup> Bezogen auf den 10-Jahres-Zeitraum, der dem 1-Jahres-Zeitraum voranging («vergangen»).

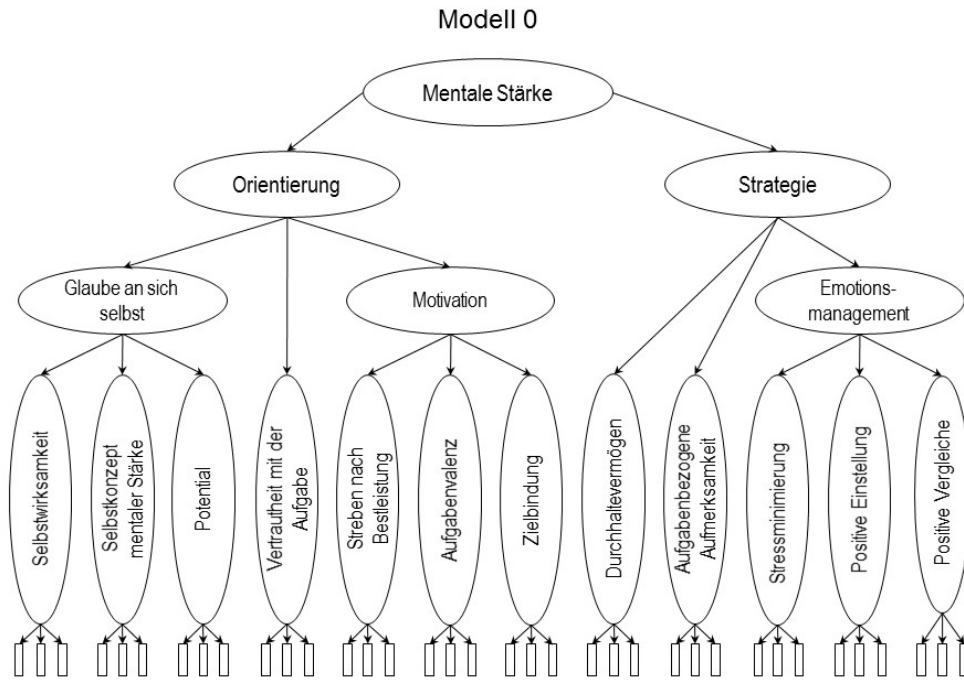
$r_{krit}$  (5 %, *n* = 93, einseitig) = .17;  $r_{krit}$  (1 %, *n* = 93, einseitig) = .24;  $r_{krit}$  (5 %, *n* = 115, zweiseitig) = .19;  $r_{krit}$  (1 %, *n* = 115, zweiseitig) = .24;  $r_{krit}$  (5 %, *n* = 115, einseitig) = .16;  $r_{krit}$  (1 %, *n* = 115, einseitig) = .22.

*Tabelle 1.* Deutsch- und englischsprachige Skalenbezeichnungen und Beispielitems des MTI-D

Skala	Beispiel
Zielbindung (ZB) – Goal Commitment	Was auch immer passiert, ich setze mich für meine Ziele ein.
Streben nach Bestleistung (SNB) – Personal Best	Mein Allerbestes zu geben, ist alles, worum es geht.
Stressminimierung (SM) – Stress Minimisation	Ich kann die Auswirkungen von Stress gut gering halten.
Selbstwirksamkeit (SW) – Self-Efficacy	Egal wie groß der Druck ist, ich glaube weiter an mich.
Aufgabenbezogene Aufmerksamkeit (AA) – Task-Specific Attention	Ich konzentriere mich voll und ganz auf die Aufgabe, nichts lenkt mich ab.
Potential (PO) – Potential	Ich habe großes Potential in diesem Bereich und werde es auch ausschöpfen.
Durchhaltevermögen (DV) – Perseverance	Ich arbeite so lange an etwas, bis ich es geschafft habe.
Positive Vergleiche (PV) – Positive Comparison	Zu sehen, dass meine Konkurrenten nicht gut zurechtkommen, erhöht meine Zuversicht.
Aufgabenvalenz (AV) – Value	Ich bewerte diesen Bereich als etwas vom Wichtigsten in meinem Leben.
Vertrautheit mit der Aufgabe (VA) – Task Familiarity	Meine Erfahrung in diesem Bereich macht mich in einer Leistungssituation stärker.
Positive Einstellung (PE) – Positivity	Wenn etwas schlecht ist, versuche ich es in etwas Positives umzudeuten.
Selbstkonzept mentaler Stärke (SK) – Mental Self-Concept	Ich weiß, dass ich über große mentale Stärke verfüge.

*Anmerkungen:* Testinstruktion (Sie-Form): Dieser Fragebogen dient der Erhebung der mentalen Stärke einer Sportlerin bzw. eines Sportlers. Es gibt keine richtigen oder falschen Antworten. Bitte kreuzen Sie bei jeder Aussage nur eine Antwort an und beantworten Sie alle Fragen. Beantworten Sie diese Fragen in Bezug auf Ihre spezifische Sportart. Wird in den Aussagen von *Bereich* gesprochen, ist damit Ihre spezifische Sportart gemeint.

Die vollständige Version des MTI-D wie auch des MTI kann bezogen werden bei Dr. Cory Middleton, Talent Development Academy, Sydney, Australia, E-Mail: [cory@tdacademy.com.au](mailto:cory@tdacademy.com.au), Phone +61 409 905 152.



*Abbildung 1.* Graphische Darstellung unterschiedlicher MTI-D-Modelle. F = Faktor erster Ordnung. HF = Faktor zweiter Ordnung. G = Globaler Faktor und S = Spezifischer Faktor. Zwecks grösserer Klarheit sind in den Modellen A, B, C und D nicht alle Faktoren (F, S) wiedergegeben. Rechtecke symbolisieren Indikatorvariablen, einseitige Pfeile Faktorladungen und zweiseitige Pfeile Kovarianzen zwischen Faktoren. Nicht dargestellt sind Residualvarianzen.

### *Abklärungen zur Retest-Reliabilität der Skalen des MTI-D*

Die Reliabilität der MTI-D-Skalen wurde anhand der Methode der Konsistenzanalyse und der Testwiederholung bestimmt. Die Ausführungen zur Konsistenzanalyse sind im Manuskript enthalten, während diejenigen zur Retest-Reliabilität in diesem Elektronischen Supplement zusammengetragen sind.

#### *Methode*

Die Retest-Reliabilität bzw. zeitliche Stabilität wurde anhand einer der beiden Jahreskohorten von Personen geschätzt, die im ersten Semester Sportwissenschaft an der Universität Bern studieren (Vollerhebung) und die nach der Erstmessung ( $t_1$ ) im Abstand von 5 ( $t_2$ ) und 7 Monaten ( $t_3$ ) nochmals befragt wurden. Von den 115 Sportstudierenden dieser Teilstichprobe, die sich an der Erstmessung ( $t_1$ ) beteiligt hatten, konnten jedoch nicht alle für die Zweit- oder Drittmessung gewonnen werden (Ausfall  $t_2$ : 41, Ausfall  $t_3$ : 33). Da der Stichprobenausfall weder mit dem Geschlecht, noch dem Alter, noch dem Wettkampf- und Leistungsniveau signifikant zusammenhing, wurden Complete-Case-Analysen durchgeführt.

Zur Schätzung der Reliabilität bzw. Stabilität für das 2-, 5- und 7-Monate-Intervall ( $t_2$ - $t_3$ ,  $t_1$ - $t_2$ ,  $t_1$ - $t_3$ ) wurde angesichts der kleinen Stichprobe die Rangkorrelation ( $\rho$ ) verwendet. Ferner wurden mittels GEE-Modellen (Generalized Estimating Equations, deutsch: generalisierte Schätzgleichungen) die Tracking-Koeffizienten berechnet (Twisk, 2013). Dabei handelt es sich um den standardisierten Regressionskoeffizienten ( $\beta_s$ ), der sich ergibt, wenn der Messwert einer Skala zum Zeitpunkt  $t_1$  als Prädiktor für die weitere Messwertentwicklung auf dieser Skala in die Regression eingeführt wird. Der Koeffizient  $\beta_s$  lässt sich als längsschnittliche Korrelation deuten und eignet sich speziell für den vorliegenden Fall unbalancierter Längsschnittdaten (ungleiche Zeitintervalle, interindividuell unterschiedliche Anzahl Messungen).

#### *Ergebnisse und Diskussion*

Die Befunde zur Retest-Reliabilität bzw. zeitlichen Stabilität des MTI-D über zwei, fünf und sieben Monate finden sich in Tabelle 3 dieses Elektronischen Supplements. Die Korrelation der Gesamtskala für die drei Intervalle betrug .70, .74 und .78; die entsprechenden gemittelten Korrelationen der Subskalen beliefen sich auf .66, .68 und .68.

Betrachtet man die längsschnittliche Korrelation, ergab sich für den Tracking-Koeffizienten ( $\beta_s$ ) der Gesamtskala ein Wert von .76; im Falle der Subskalen lag er sich zwischen .58 (*Vertrautheit mit der Aufgabe*) und .78 (*Zielbindung*).

Leider liegen zum Original keine Retest-Koeffizienten vor und auch keine Überlegungen zur Merkmalfluktuation. Deshalb sind diese Befunde für das MTI-D nicht gut einzuordnen. Insgesamt deuten sie aber auf relativ messgenaue Skalen oder ein relativ stabiles Konstrukt hin.

*Tabelle 3. Retest-Reliabilität des MTI-D*

MTI-D-Skalen	5 Monate ( $t_1-t_2$ )	7 Monate ( $t_1-t_3$ )	2 Monate ( $t_2-t_3$ )	<i>M</i>	Tracking- Koeffizient ( $\beta_s$ )
Globalfaktor (G)/Gesamtskala	.74	.78	.70	.74	.76
Zielbindung (ZB)	.79	.77	.79	.78	.78
Streben nach Bestleistung (SNB)	.56	.63	.55	.58	.58
Stressminimierung (SM)	.68	.74	.65	.69	.70
Selbstwirksamkeit (SW)	.61	.65	.65	.64	.61
Aufgabenbezogene Aufmerksamkeit (AA)	.74	.70	.69	.71	.72
Potential (PO)	.74	.73	.75	.74	.74
Durchhaltevermögen (DV)	.60	.67	.69	.65	.64
Positive Vergleiche (PV)	.63	.69	.63	.65	.65
Aufgabenvaleanz (AV)	.65	.61	.71	.66	.65
Vertrautheit mit der Aufgabe (VA)	.55	.58	.61	.58	.56
Positive Einstellung (PE)	.60	.74	.77	.68	.66
Selbstkonzept mentaler Stärke (SK)	.70	.59	.64	.65	.63
<i>M</i> (Subskalen)	.66	.68	.68	.67	.66

*Anmerkungen.* Rangkorrelationen nach Spearman ( $\rho$ ). Die Analysen zur Retest-Reliabilität beruhen auf einer Teilstichprobe der Studierenden der Sportwissenschaft (5 Monate:  $n = 74$ ; 7 Monate:  $n = 82$ ; 2 Monate:  $n = 64$ ). Alle  $\beta_s$  sind signifikant ( $\chi^2$ -Test,  $df = 1$ ,  $p < .001$ ).

## Literatur

Twisk, J. W. R. (2013). *Applied longitudinal data analysis for epidemiology: A practical guide* (2nd ed.). Cambridge: Cambridge University Press.



Tabelle 3. Standardisierte Faktorladungen ( $\lambda$ ), Residualvarianzen ( $\varepsilon$ ) und Reliabilitätskoeffizienten (McDonalds  $\omega$ ) des BESEM-Modells D ( $N = 1\ 122$ )

Items	MTI-D-Faktoren														Nebenladungen	
	G	S1	S2	S3	S4	S5	S6	S7	S8	S9	S10	S11	S12	$\varepsilon$	$M$	Maximum
	Global	ZB	SNB	SM	SW	AA	PO	DV	PV	AV	VA	PE	SK			
Item 1	.68	<b>.44</b>	-.04	-.15	-.15	-.02	-.07	.06	.01	.03	.01	-.12	-.16	.24	.07	.16
Item 13	.79	<b>.14</b>	.01	<b>-.21</b>	-.19	-.03	-.10	.00	-.09	.00	-.06	-.14	-.17	.20	.09	.21
Item 25	.76	<b>.08</b>	.10	-.17	-.03	.04	-.07	.10	-.05	.00	-.09	-.04	-.06	.34	.07	.17
Item 2	.64	<b>.68</b>	<b>.09</b>	-.05	.01	.01	-.12	.02	.01	.03	.04	-.06	-.10	.10	.10	.68
Item 14	.62	.11	<b>.45</b>	-.06	-.07	-.01	.00	.07	.00	.14	-.03	-.03	-.11	.35	.06	.14
Item 26	.64	.15	<b>.49</b>	-.07	-.01	-.03	-.12	.05	.00	.13	-.01	-.04	-.08	.29	.06	.15
Item 3	.48	-.03	-.06	<b>.62</b>	.00	.04	-.08	-.03	-.02	-.16	-.03	-.01	.08	.34	.05	.16
Item 15	.55	-.07	.05	<b>.34</b>	.01	.03	-.03	-.06	.06	-.07	-.04	.17	-.01	.54	.09	.17
Item 27	.61	-.11	-.01	<b>.53</b>	.06	.07	-.05	-.05	-.01	-.11	-.08	.07	.08	.29	.06	.11
Item 4	.66	.02	-.16	.15	<b>.19</b>	-.02	-.03	-.14	-.02	-.16	-.07	.08	.12	.41	.09	.16
Item 16	.74	-.04	-.09	.12	<b>.27</b>	-.01	-.04	-.09	-.05	<b>-.21</b>	-.09	.03	.06	.29	.08	.21
Item 28	.75	-.09	-.06	.07	<b>.44</b>	-.01	.02	-.05	-.03	-.17	-.09	.02	.11	.17	.06	.17
Item 5	.64	.09	-.05	.06	-.06	<b>.50</b>	-.07	-.02	-.02	-.09	-.06	.02	-.01	.31	.05	.09
Item 17	.68	-.08	.03	.07	-.02	<b>.51</b>	-.02	.05	-.04	-.02	-.05	-.02	-.01	.26	.04	.08
Item 29	.71	.01	-.02	.03	.08	<b>.48</b>	.01	.03	-.01	-.06	-.01	.04	.05	.24	.03	.08
Item 6	.48	.00	-.07	-.03	-.03	-.04	<b>.66</b>	-.01	.02	.10	.03	-.02	-.03	.31	.03	.10
Item 18	.60	-.03	.03	-.07	.01	-.01	<b>.59</b>	.00	.03	.11	.00	-.05	.03	.28	.03	.11
Item 30	.52	-.11	.00	-.02	.03	.01	<b>.70</b>	.07	.01	.15	.08	-.02	-.01	.19	.05	.15
Item 7	.60	.04	-.04	.06	-.09	.09	.06	<b>.08</b>	.05	.03	-.08	.07	-.16	.57	.07	.16
Item 19	.67	.03	.04	-.07	-.12	.00	.08	<b>.29</b>	-.03	.07	-.03	-.02	-.07	.43	.05	.12
Item 31	.73	.05	.05	-.08	-.02	.03	.09	<b>.62</b>	-.01	.03	.02	-.01	-.06	.07	.04	.09
Item 8	.31	.01	-.02	.04	-.05	.03	.00	-.01	<b>.79</b>	.03	-.02	-.05	-.06	.27	.10	.06
Item 20	.39	.03	.00	-.04	.04	-.03	.01	-.04	<b>.77</b>	.04	.00	.02	.01	.25	.09	.04
Item 32	.35	-.02	.02	-.01	.01	-.01	.01	.03	<b>.86</b>	.04	.03	.00	.03	.14	.09	.04

Items	MTI-D-Faktoren														Nebenladungen	
	G	S1	S2	S3	S4	S5	S6	S7	S8	S9	S10	S11	S12	$\varepsilon$	$M$	Maximum
	Global	ZB	SNB	SM	SW	AA	PO	DV	PV	AV	VA	PE	SK			
Item 9	.58	.09	.15	-.09	-.03	.01	.16	.00	.07	<b>.49</b>	.00	-.04	-.09	.35	.07	.16
Item 21	.52	-.01	.02	-.05	-.02	-.01	.09	.01	.07	<b>.78</b>	.06	-.04	-.04	.09	.04	.09
Item 33	.56	.02	-.01	-.06	-.01	-.06	.13	.00	.04	<b>.67</b>	.12	-.06	-.05	.18	.05	.13
Item 10	.60	.09	.00	.02	-.12	-.03	.13	-.11	.03	.02	<b>.09</b>	-.04	.02	.58	.06	.13
Item 22	.65	.11	.01	-.05	-.09	-.07	.00	-.02	.01	.18	<b>.31</b>	.01	-.03	.42	.05	.18
Item 34	.62	.03	-.01	-.04	-.02	-.03	.08	.01	.02	.15	<b>.76</b>	-.02	-.02	.00	.04	.15
Item 11	.57	-.11	-.07	.08	-.03	.02	-.05	-.03	-.06	-.14	-.04	<b>.56</b>	.06	.31	.06	.14
Item 23	.62	-.06	.03	-.01	.00	.01	-.06	.01	-.02	-.07	-.03	<b>.58</b>	.06	.27	.03	.07
Item 35	.69	.04	-.02	.07	.17	-.01	-.02	.03	-.06	-.12	.04	<b>.39</b>	.08	.31	.06	.17
Item 12	.67	.01	-.07	<b>.25</b>	.01	-.04	<b>-.20</b>	-.11	-.04	-.21	-.06	.13	<b>.25</b>	.31	.10	.25
Item 24	.63	-.17	-.06	.09	.02	.03	.01	-.05	-.01	-.13	-.03	.09	<b>.56</b>	.22	.06	.17
Item 36	.68	-.10	-.05	.12	.05	-.01	-.04	-.04	-.03	<b>-.21</b>	-.04	.07	<b>.52</b>	.19	.07	.21
$M$ (Hauptladungen)	.62	.23	.35	.51	.30	.50	.65	.35	.81	.66	.44	.51	.45			
$M$ (Nebenladungen)		.08	.04	.08	.05	.03	.06	.04	.03	.10	.04	.05	.06			
Maximum (Nebenladungen)		.68	.16	.25	.19	.09	.20	.14	.09	.21	.12	.17	.17			
omega ( $\omega$ ) bzw. omegaS ( $\omega_S$ )	.98	.87	.86	.81	.86	.89	.89	.82	.91	.91	.83	.87	.89			
omegaH ( $\omega_H$ ) bzw. omegaHS ( $\omega_{HS}$ )	.94	.07	.20	.37	.13	.31	.53	.19	.77	.53	.23	.35	.28			
$\omega_H/\omega$ bzw. $\omega_{HS}/\omega_S$	.96	.08	.23	.45	.15	.35	.60	.24	.84	.58	.28	.40	.31			

Anmerkungen: BESEM = Bifactor exploratory structural equation modeling. G = Globalfaktor; S = spezifischer Faktor; ZB = Zielbindung; SNB = Streben nach Bestleistung; SM = Stressminimierung; SW = Selbstwirksamkeit; AA = Aufgabenbezogene Aufmerksamkeit; PO = Potential; DV = Durchhaltevermögen; PV = Positive Vergleiche; AV = Aufgabenvaleanz; VA = Vertrautheit mit der Aufgabe; PE = Positive Einstellung; SK = Selbstkonzept mentaler Stärke. Fett hervorgehoben sind die Ladungen auf die angezielten Faktoren sowie Ladungen auf nicht angezielte Faktoren  $\geq .20$ . omegaS,  $\omega_S$  = omega subscale; omegaH,  $\omega_H$  = omega hierarchical; omegaHS,  $\omega_{HS}$  = omega hierarchical subscale (Rodriguez, Reise & Haviland, 2016).

*Tabelle 4. Korrelationen zwischen MTI-D-Faktorwerten und MTI-D-Skalen (manifeste Ebene, N = 1 122)*

MTI-D-Faktorwerte	MTI-D-Skalen												
	Gesamt- skala	ZB	SNB	SM	SW	AA	PO	DV	PV	AV	VA	PE	SK
Globalfaktor (G)	.98	.86	.75	.66	.82	.77	.60	.81	.39	.61	.76	.72	.75
Zielbindung (ZB)	.05	.28	.38	-.09	-.05	.01	-.06	.06	.01	.04	.10	-.06	-.11
Streben nach Bestleistung (SNB)	.04	.04	.54	-.01	-.15	-.02	-.02	.03	.00	.07	.00	-.03	-.09
Stressminimierung (SM)	.07	-.24	-.08	.71	.15	.08	-.05	-.04	.00	-.09	-.04	.06	.20
Selbstwirksamkeit (SW)	.02	-.17	-.04	.03	.43	.00	.00	-.12	.00	-.02	-.11	.06	.04
Aufgabenbezogene Aufmerksamkeit (AA)	.07	-.01	-.01	.07	-.02	.66	-.02	.06	-.01	-.03	-.07	.01	-.01
Potential (PO)	.10	-.10	-.10	-.07	-.02	-.03	.80	.10	.01	.15	.09	-.06	-.09
Durchhaltevermögen (DV)	.03	.06	.06	-.06	-.11	.03	.03	.43	-.01	.00	-.05	.00	-.08
Positive Vergleiche (PV)	.13	-.05	.01	.01	-.04	-.03	.02	.01	.92	.07	.03	-.05	-.03
Aufgabenvalenz (AV)	.08	.02	.13	-.14	-.22	-.07	.14	.06	.04	.76	.15	-.13	-.22
Vertrautheit mit der Aufgabe (VA)	.03	-.05	.00	-.06	-.10	-.04	.04	-.04	.00	.07	.48	-.01	-.05
Positive Einstellung (PE)	.08	-.13	-.06	.11	.06	.02	-.03	.02	-.01	-.06	-.02	.69	.13
Selbstkonzept mentaler Stärke (SK)	.05	-.18	-.14	.07	.13	.01	-.01	-.14	-.01	-.08	-.01	.09	.60

*Anmerkungen:* Die durchschnittliche Korrelation der MTI-D-Faktorwerte und MTI-D-Skalen beträgt .70.  $r_{\text{krit}}(5\%, n = 1\,122, \text{einseitig}) = .05$ ;  $r_{\text{krit}}(1\%, n = 1\,122, \text{einseitig}) = .07$ .