

Developing and training mental toughness in sport: a systematic review and meta-analysis of observational studies and pre-test and post-test experiments

Andreas Stamatis ¹, Peter Grandjean,² Grant Morgan,³ Robert Noah Padgett,³ Richard Cowden,⁴ Panagiotis Koutakis⁵

To cite: Stamatis A, Grandjean P, Morgan G, *et al*. Developing and training mental toughness in sport: a systematic review and meta-analysis of observational studies and pre-test and post-test experiments. *BMJ Open Sport & Exercise Medicine* 2020;**6**:e000747. doi:10.1136/bmjsem-2020-000747

► Additional material is published online only. To view please visit the journal online (<http://dx.doi.org/10.1136/bmjsem-2020-000747>).

Accepted 19 May 2020



© Author(s) (or their employer(s)) 2020. Re-use permitted under CC BY-NC. No commercial re-use. See rights and permissions. Published by BMJ.

For numbered affiliations see end of article.

Correspondence to

Dr Andreas Stamatis;
coach_stam@rocketmail.com

ABSTRACT

Objective To investigate the efficacy of interventions designed to train and develop mental toughness (MT) in sport.

Design Systematic review and meta-analysis.

Data sources Journal articles, conference papers and doctoral theses indexed in Embase, Scopus, PubMed and SPORTDiscus from inception to 22 November 2019.

Eligibility criteria for selecting studies

Observational and pre–post experimental designs on the efficacy of physical and/or psychological interventions designed to promote MT in athletes.

Results A total of 12 studies, published between 2005 and 2019, were included in the review. A majority of the studies included a sample comprised exclusively of male athletes (54.55%), MT interventions were primarily psychological (83.33%) and most studies measured MT via self-report (75%). The Psychological Performance Inventory (25%), the Mental Toughness Questionnaire-48 (16.67%), and the Mental, Emotional and Bodily Toughness Inventory (16.67%) were the most popular inventories used to measure MT. Methodological quality assessments for controlled intervention studies ($k=7$), single group pre–test–post–test designs ($k=4$) and single-subject designs ($k=1$) indicated that the risk of bias was high in most (75%) of the studies. The meta-analysis involving $k=10$ studies revealed a large effect ($d=0.80$, 95% CI 0.30 to 1.28), with variability across studies estimated at 0.56.

Conclusion Although the findings of this review suggest there are effective, empirically based interventions designed to train MT in sport, practitioners should be aware of the level of validity of intervention research before adopting any of the MT training programmes reported in the applied sport psychology literature.

INTRODUCTION

Mental toughness (MT) has been linked to a variety of desirable behaviours (eg, perseverance) and favourable sport performance outcomes (eg, achievement), making it highly attractive to various stakeholders involved in sport.^{1 2} However, the practical utility of

Summary box

- An abundance of qualitative research suggests that mental toughness (MT) may be cultivated by manipulating features of the sporting environment, yet practitioners seem restricted by a lack of evidence on best practices for developing MT.
- Results from this systematic review and meta-analysis of observational studies and pre-test and post-test experimental designs revealed a large effect for interventions designed to train and develop MT in sport.
- Included studies were found to have a moderate or high risk of bias and there were several methodological limitations identified that affect the interpretability (eg, insufficient reporting of methodology) and generalisability (eg, under-representation of female athletes) of the reported findings.
- Practitioners should be aware of the level of validity of MT intervention research before adopting any of the MT training programmes reported in the applied sport psychology literature.

MT for anyone who is looking to successfully navigate the multifaceted demands of being a competitive athlete hinges on scientific efforts dedicated towards rigorous development of efficacious evidence-based MT interventions.³ After several years of accumulating research in which scholars have used knowledge acquired on processes involved in MT development to design and empirically test MT interventions, the purpose of this study was to assess the efficacy of MT interventions through a systematic review of existing evidence and an accompanying meta-analysis.

Overview of MT

Over the last 20 years, considerable progress has been made towards resolving historical concerns about the conceptual ambiguity (for an overview, see³) of MT, its definition

(for a chronology of primary MT definitions, see⁴) and its measurement (for a summary of existing MT instruments, see⁵). Most prominent conceptual models share cognitive (eg, generalised self-efficacy), emotional (eg, regulation of emotions) and behavioural (eg, perseverance) features thought to be central to MT. Scholars generally agree that MT is a psychological resource that enables athletes to initiate and sustain goal-oriented efforts towards achieving optimal levels of performance in the face of stressors that vary in duration, frequency and intensity (eg,⁶⁻⁸). However, there remain several areas of conceptual ambiguity surrounding MT (for additional details, see^{4,9}), which highlight some of the challenges sport personnel (eg, coaches, trainers, sport psychology practitioners) are likely to encounter as they evaluate the quality of MT interventions that have been tested and make decisions about which approaches are suitable for targeting the MT of athletes they work with.

MT development in sport

A number of qualitative studies have explored the experiences of athletes, coaches and sport psychologists to understand the developmental process of MT (for a review, see¹⁰). Although there have recently been reservations about whether targeted interventions to improve MT are merited (eg,¹¹), most findings have revealed a complex, long-term trajectory of MT development in athletes that may be cultivated independently or conjunctively by features of the sporting environment (eg, motivational climate) and experiences that occur outside of sport (eg, critical life events).

Given the potential to promote MT by adapting sporting environments (eg,¹²), controllable features of the sport context (eg, training conditions) are of particular interest to many stakeholders (eg,¹³). Nonetheless, practitioners seem restricted by the lack of evidence-based information on best practices for developing and maintaining MT.^{14 15} In light of recent events where training programmes designed to test and train MT have resulted in hospitalisation or death of elite-level athletes,^{16 17} progress towards evidence-based MT training practices requires an improved understanding of effective strategies for promoting MT. Thus, it appears necessary to take stock of and comprehensively evaluate the efficacy of empirically tested MT interventions that have emerged over the last two decades.

Statement of purpose of the present study

Several narrative and systematic reviews addressing different aspects of MT (eg, developmental processes, success in sport) have been published (eg,^{1 4 10 11}). Following the PRISMA guidelines (see¹⁸), the current review builds on the existing literature in two main ways. First, there has been no attempt to systematically collate all available empirical evidence on the efficacy of MT training in sport from observational and pre-test and post-test experimental studies published up to 2019. The PRISMA guidelines provide a framework for

systematically evaluating the available evidence on MT interventions in sport, which is advantageous because it eliminates bias¹⁹ and is a useful reporting approach for practitioners interested in using the findings to inform their intervention work.²⁰ Second, a meta-analysis is performed to summarise effect sizes across studies and generate more precise estimates of intervention effects by combining estimates reported in individual studies.²¹ Quantitative, qualitative and mixed-methods observational and experimental studies that have been conducted on MT interventions in sport are reviewed and, where available, effect sizes at post-test are summarised. Thus, the purpose of this study is to systematically review and meta-analyse existing evidence to address the following research question: Are there any effective, empirically based interventions designed to train MT in sport?

METHOD

Protocol and registration

Methods of the systematic review process and meta-analysis were determined and described in advance. They were documented in a review protocol (see online supplementary material.). The protocol pre-determined the purposes and processes of the systematic review, thereby limiting the probability of bias.²² Although this systematic review was not registered, the protocol was not modified during the research process.

Eligibility criteria

Types of studies

The study protocols included quantitative, qualitative and mixed-methods observational studies and pre-test and post-test experiments studying the effect of MT interventions on MT levels of athletes. Case studies were excluded. No publication date restriction was imposed, but the studies were limited to those written in English.

Types of participants

Participants included those of any age or gender involved in a sport where they received any level of MT intervention (for the purposes of this systematic review, the term 'MT intervention' was applied broadly to include studies that used any type of psychological and/or physical approach designed specifically to target MT in athletes, as well as studies that evaluated changes in MT as a byproduct of athletic training protocols that were not necessarily targeting MT). MT training was defined as any type of psychological and/or physical strategies that were prescribed to athletes aiming to influence their MT levels in relation to their sport.

Types of interventions

Studies that address physical and/or psychological interventions targeting MT in sport were included in this review.

Types of outcome measures

Self-reported and other-reported MT (eg, athletes assessed by their coaches) scores was the primary outcome

measure of treatment effect. MT scores were acquired through various measures of MT (includes instruments developed specifically to measure MT, as well as other instruments that were used by researchers to make inferences about athletes' MT (eg,²³)).

Information sources

The authors identified studies by searching the following electronic databases: Embase, Scopus, PubMed and SPORTDiscus. The last search was run on 22 November 2019.

Search

This study can be repeated or updated through the use of the following search terms: (1) Embase: 'mental toughness' OR 'mentally tough' AND ('sport'/exp OR sport OR 'sports') AND english/lim; (2) Scopus: ("mental* tough*" AND sport*) AND (LIMIT-TO (DOCTYPE, "ar") OR LIMIT-TO (DOCTYPE, "re") OR LIMIT-TO (DOCTYPE, "ip")) AND (LIMIT-TO (Language, "English")); (3) PubMed: ((Mental* AND Tough* [Title/Abstract] AND sport* [Title/Abstract])) AND english[Language]; (4) SPORTDiscus: TI ("mental* tough*" AND sport*) OR AB ("mental* tough*" AND sport*) OR ("mental* tough*" AND sport*).

Study selection

Two reviewers were involved in the process of study selection. After identifying the keywords and the databases, one of the reviewers performed the initial search (ie, identification) and screening (see figure 1). Next, both reviewers re-examined the process up to that point, which included an evaluation of the initial set of records that were identified after screening. No additional searches were performed. Lastly, based on the existing recommendations of screening records,²⁴ both reviewers performed an eligibility assessment independently. There were no disagreements.

Data collection process

A data extraction sheet was developed, but it was not pilot tested. Both reviewers extracted the data from the included studies (see Data items). Disagreements about extracted data were resolved by discussion between the reviewers until consensus was reached. Lead authors of two selected studies were contacted to acquire additional numerical data that were needed. The requested information was not provided, leading to the exclusion of those two studies from the meta-analysis (ie,^{23 25}).

Data items

The authors extracted the following information from each of the included studies: (1) general characteristics of the publication: title, author/s, year of publication, institution of corresponding author/s, country of corresponding author/s, name of journal and purpose of the study; (2) characteristics of the participants: sport type, sport level, type of participants (eg, athletes, coaches), and quantity, age and sex of athlete participants; (3)

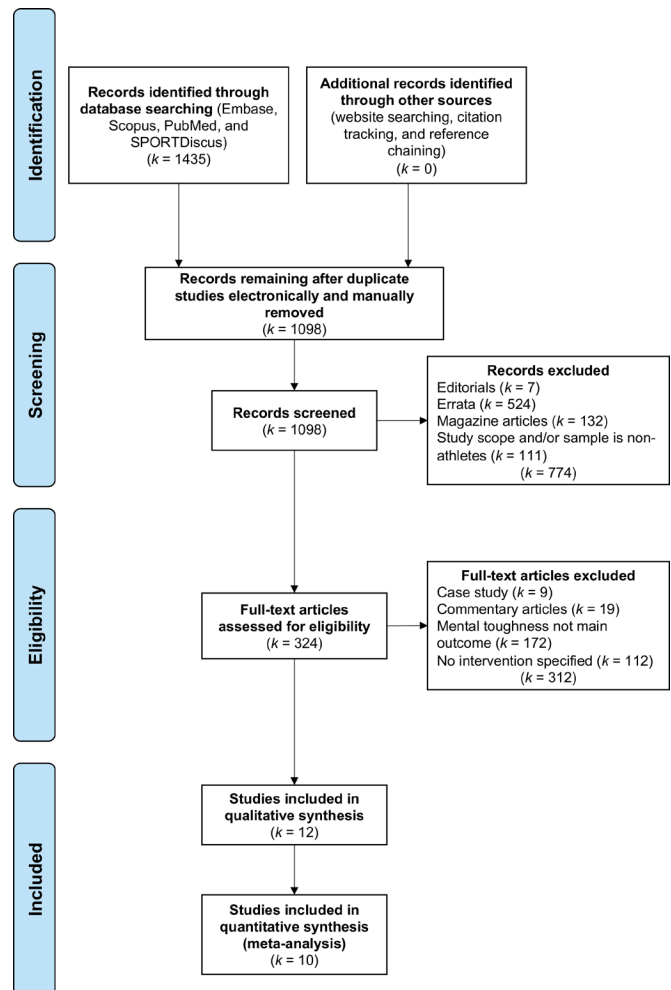


Figure 1 Flow of information through the different phases of the systematic review.

intervention—type and duration; (4) research strategy—overall research design, number of experimental groups and number of participants per group; (5) instruments—type of MT instrument used and rater/informant of MT (eg, athlete, coach); (6) analysis and results—primary statistical analysis, main outcomes and variables included (see online supplementary table S1).

Risk of bias in individual studies

Controlled intervention studies

The two reviewers worked independently and used the Physiotherapy Evidence Database (PEDro) scale to ascertain the validity of the seven controlled intervention studies (CISs), three of which were randomised controlled trials (RCTs; for details on this approach, see²⁶). Reviewers evaluated those seven studies against the PEDro criteria using a dichotomous scale ($N=0$; $Y=1$). With the exception of the first criterion (ie, eligibility criteria were specified), which affects external rather than internal validity, total quality scores were created by summing the number of affirmative answers to the items on the scale (see online supplementary table S2). The total score (range=0–10) reflects the internal validity of the study and whether it

contains sufficient statistical information to make it interpretable. A higher score is representative of a study with higher methodological quality. Studies with scores of ≥ 6 have a low risk of bias.²⁷

Single-group pre-test–post-test designs

The two reviewers independently assessed the four studies meeting criteria for inclusion using the ‘Quality Assessment Tool for Before–After (Pre–Post) Studies With No Control Group’.²⁸ Reviewers rated 12 items designed to assess the quality of each study using a categorical rating approach (see online supplementary table S3a). Thereafter, reviewers evaluated for risk of bias and provided an overall quality rating by holistically evaluating their responses to the items for each study. Reviewers used an ordinal scale of *Good*, *Fair* and *Poor* to rate the quality of each study (see online supplementary table S3b).

Single-subject design

The two reviewers independently evaluated the one study that met criteria for inclusion using the ‘Quality Indicators’ measure (for details on this approach, see²⁹). The measure contains 28 indicators that are distributed across eight domains (one to six indicators per domain) of methodological quality (see online supplementary table S4). In this study, six of the original indicators (ie, 6.4, 6.8, 6.9, 7.6, 8.1 and 8.3) were excluded from evaluation because they apply only to group comparisons. For a study to be characterised as ‘methodologically sound’, it has to meet the standards of all the quality indicators.

Summary measures

The primary outcome of interest is self-reported and other-reported MT ratings. MT scores were translated into standardised variables. The intended summary effect measures are the standardised mean differences (SMDs), which can be calculated for both between-subjects and within-subjects designs. The SMDs were estimated in accordance with the design of each study. For a discussion of the estimation of SMDs for between-subjects and within-subject designs, see,³⁰ and for the estimation of SMDs for single case designs, see.³¹ The SMDs, along with their CIs and sampling variances, were calculated using the ‘escalc’ command of the ‘metafor’ package³² in R.³³

Planned methods of analysis

The meta-analysis was completed using a random-effects model. As such, the model treats each effect size estimate as having been sampled at random from a universe of possible effects, and the estimates will vary between studies. The model estimates this random effect as a variance component to provide an estimate of how much MT intervention effects might be expected to vary between studies. The model also includes a fixed-effect estimate of the MT intervention effect, which is the average standardised effect across studies. The random-effects model can be expressed as

$$\theta_i = \mu + u_i,$$

where μ is the average effect of MT intervention and u_i is the random error term distributed $N(0, \tau^2)$. The random-effects model parameters were estimated using the ‘rma.mv’ command of the ‘metafor’ package in R with the restricted maximum-likelihood estimator. The random-effect and fixed-effect estimates are reported along with their associated 95% CIs.

Risk of bias across studies

Except for bias related to search terms and databases used, we acknowledge that searching databases does not ensure all relevant studies were discovered during the search process. For example, unpublished technical reports and master’s theses were likely under-represented. In addition, publication bias towards positive results may have had an influence on the inclusion of articles in the meta-analysis. The potential for publication bias was investigated using a funnel plot, which graphically represents the relationship between observed effect sizes and estimation precision (ie, standard errors). Ideally, the plot appears symmetrical and points on the plot form an inverted funnel. Studies with smaller sizes tend to be scattered widely at the bottom of the plot. However, studies with larger sample sizes typically have greater precision in the estimates of effected sizes and are centrally located at the top of the funnel. Asymmetrical plots or blank spots within plots (eg, lower right-hand corner) are evidence of potential bias.

Additional analyses

Due to the small number of articles that met the criteria for inclusion in the meta-analysis, the potential effects of moderator variables were not examined in this study.

RESULTS

Study selection

The systematic review consisted of four phases (see figure 1). The first phase (Identification) resulted in 1435 results. The second phase (Screening) entailed automatically and manually removing duplicate records, reducing the total to 324 articles. In the third phase (Eligibility), the remaining articles were checked via the eligibility criteria determined in the Protocol (eg, English only). As a result, 312 articles were excluded. All of the 12 studies^{23 25 34–43} (articles⁴⁴ and⁴⁰ were both part of a larger mixed-methods project that resulted in two separate publications, one involving the qualitative component the other the quantitative component; the qualitative study⁴⁴ did not adequately fulfil the inclusion criteria on its own and was therefore excluded from the review) that entered the fourth phase (Included) were publications in peer-reviewed journals. Two studies for which requested information from lead authors could not be acquired (ie,^{23 25}) were excluded from the meta-analysis.

Study characteristics

The most important and relevant-to-the-review-question characteristics for each of the 12 studies included in

Methodological and statistical rigor	Research design	Key areas that increased risk of bias	Common limitations affecting bias and generalisability	Methodological and reporting recommendations for future research
More rigorous	Controlled intervention studies	<ul style="list-style-type: none"> Unspecified eligibility criteria. Non-random assignment. Allocation of participants not concealed. Subjects not blinded. Therapists/facilitators not blinded. Assessors not blinded. ≤ 85% response rate at follow-up. 	<ul style="list-style-type: none"> 83.33% of studies used a single rating source to measure MT; 75% of studies used self-report alone to measure MT. A wide range of measures were used to assess MT across studies, which vary in theoretical grounding, type and dimensions of MT assessed, and psychometric quality. 	<ul style="list-style-type: none"> Use of well-designed and rigorously conducted randomised controlled trials to reduce bias and improve the quality of evidence on the efficacy of MT interventions. Design and evaluation of culturally-informed interventions that are sensitive to the broad social context and sport-specific features that influence local understandings of MT; comparisons of culturally-informed or adapted interventions to generic approaches targeting MT development.
	Single group pretest-posttest designs	<ul style="list-style-type: none"> < 100% enrollment of all subjects who met inclusion criteria. Assessors not blinded. < 80% response rate at follow-up. Absence of multiple measurements before and after treatment. 	<ul style="list-style-type: none"> 54.55% of studies recruited exclusively male samples; a single study recruited exclusively female athletes. 91.67% of studies included athletes from a single sport. 33.33% of studies reported sample sizes of ≤ 20 ($M = 48.42$, $SD = 46.94$). 91.67% of studies included samples of young people (i.e., 10 to 24 years of age); 58.33% of studies included samples comprised exclusively of adolescents (i.e., 10 to 19 years of age). 	<ul style="list-style-type: none"> Design and evaluation of theoretically-grounded physical interventions to develop MT; comparisons of interventions targeting physical, psychological, or both domains. Further development and evaluation of gradual, long-term approaches to improve MT, and comparing the efficacy of such to brief, intensive programs. Long-term follow-up assessment of MT and related outcomes (e.g., athletic performance) to evaluate stability of treatment gains.
Less rigorous	Single subject designs	<ul style="list-style-type: none"> < 3 data points during baseline phase without justification by authors. < 3 data points during treatment phase without justification by authors. 	<ul style="list-style-type: none"> 66.67% of studies consisted of athletes from Australia, Canada, the United Kingdom, and the United States of America, which are more developed, high-income countries with principally individualistic value systems. 	<ul style="list-style-type: none"> Use of multi-informant rating approaches to assess psychological and behavioral indicators of MT from sources that extend beyond self-report (e.g., coaches). Thorough reporting of sample sociodemographic (e.g., age, sex, race) and sport participation (e.g., sport type, level of participation, length of participation) characteristics. Detailed description of intervention protocol (e.g., content, procedure). Recruitment of samples with a higher representation of females.

Figure 2 Summary of methodological quality, design-specific limitations, and shared limitations affecting bias and generalisability of reviewed studies, along with recommendations for future mental toughness (MT) interventions.

this review are presented in online supplementary table S1 (eg, instrument used to collect data on MT, rater/informant of MT, type and duration of intervention). A summary of methodological trends based on extracted information is presented in figure 2.

Risk of bias within studies

Controlled intervention studies

As depicted in figure 3, final PEDro scores for most of the CISs ($k=5$) were <6 ($M=5.14$, $SD=1.36$). In general, no study satisfied all criteria (see online supplementary table S2). Few studies blinded participants ($k=1$), the therapists/facilitators administering the treatment ($k=2$) or the assessors ($k=2$). Only two studies appeared to have used allocation concealment. A small number of studies indicated that group allocation was randomised ($k=3$) or specified participant eligibility criteria ($k=4$). Similarly, most studies ($k=5$) reported a response rate of ≤85% among participants who were initially allocated to groups.

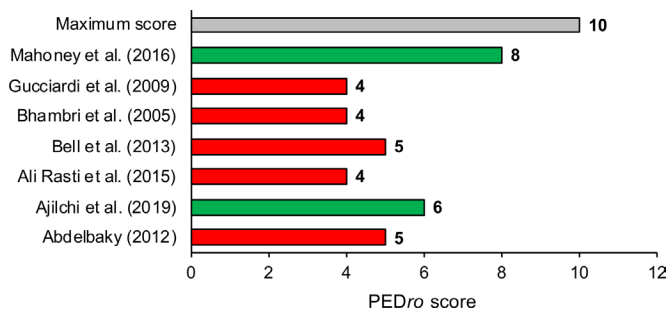


Figure 3 Total PEDro scores for $k=6$ studies.

On the other hand, in all seven studies, the results of between-group comparisons were reported. In most studies ($k=6$), all subjects received treatment or control conditions as allocated (ie, 'intention to treat') and groups were similar at baseline. A majority of studies ($k=5$) provided point measures and estimates of variability (eg, SD) for at least one primary outcome.

Single group pre-test-post-test designs

Reviewer evaluations of the four studies are presented in online supplementary table S3a (Criterion 12 of the rating instrument was not applicable to any of the studies (ie, no group-level data)) and S3b. Both raters assessed the quality of the studies as *Fair* ($k=1$) or *Poor* ($k=3$). Quality was strengthened when the research question was explicitly defined ($k=4$), eligibility criteria for selecting participants were described ($k=4$), formal statistical tests were used and levels of significance reported ($k=4$), the results could be extrapolated because participants appeared to be representative of the population from which they were drawn ($k=3$), and the intervention ($k=2$) and the outcome measures ($k=2$) were described in detail.

Concerning areas of improvement, no study blinded the outcome assessors or measured outcomes at multiple times before and after the intervention. A majority of studies did not enrol all subjects who met the inclusion criteria ($k=3$). At least half of the studies had a follow-up response rate of <80% ($k=2$) and did not provide a clear definition of the intervention or details about consistency of intervention delivery ($k=2$). There were also cases (eg, ³⁸) where relevant information was not reported

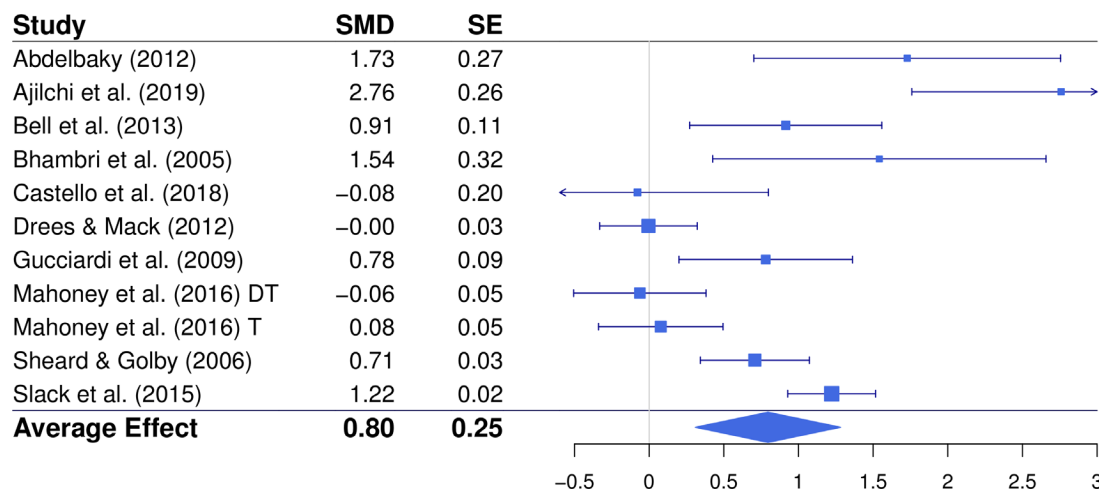


Figure 4 Summary data and estimated effects (with 95% CIs) for studies included in the quantitative analysis ($k=10$). DT, delayed treatment condition; SMD, standardised mean difference; T, treatment condition.

(eg, Item 5; statistical power) or the raters could not determine if a criterion was satisfied (eg, Item 3; sample representative of population). The design weaknesses across these studies make it difficult to determine the extent to which results can reliably be attributed to the effects of the interventions (ie, internal validity).

Single-subject design

The results of the assessment of the one single-subject study are presented in online supplementary table S4. This baseline-intervention-baseline study (ie, ABA) did not meet indicators 6.6 (ie, there were fewer than three data points in the baseline phase without justification by the authors) and 7.4 (ie, there were fewer than three data points per phase without justification by the authors). Thus, the study did not meet all of the quality indicators and could not be considered 'methodologically sound'.

Results of individual studies

Figure 4 presents the summary statistics (ie, standard mean difference and squared SE) and effect estimates for the criterion of MT in each of the included studies. Unpublished data were not included.

Synthesis of results

The estimated SMDs ranged from -0.08 to 2.76 , which are trivial to large effects. The estimated SMDs and sampling variances are shown in figure 4 (Mahoney *et al.*⁴¹ included a delayed treatment group which received the same MT intervention as the immediate treatment group after the waiting period; both groups were included in this meta-analysis to show overall effects of the MT training programme).⁴¹ The studies with the strongest effects were those with the most focused and targeted interventions for increasing MT. The average effect of MT interventions was 0.80 ($0.31 \leq \mu \leq 1.28$) SD; that is, MT scores increased by 0.80 SD following MT intervention. This is a large effect by conventional standards. The values contained within the CI were at least medium effect sizes, which lends support to the effectiveness of

MT intervention. The variance of this estimate was 0.56 ($0.18 \leq \tau^2 \leq 1.89$). The width of the CI indicates that the variability of the effect sizes may not be precise in that the variability across studies may be small (0.18) or large (1.89). This width is likely due to the small number of studies included in the meta-analysis and presence of a very large effect found in Ajilchi *et al.*³⁵ Overall, a large positive effect was found to support the efficacy of interventions for increasing MT in a variety of athlete groups.

Risk of bias across studies

The funnel plot is presented in figure 5. Deviations from the ideal distribution were difficult to identify with only 10 studies. One study might be an outlier, but not enough information was available to definitively make

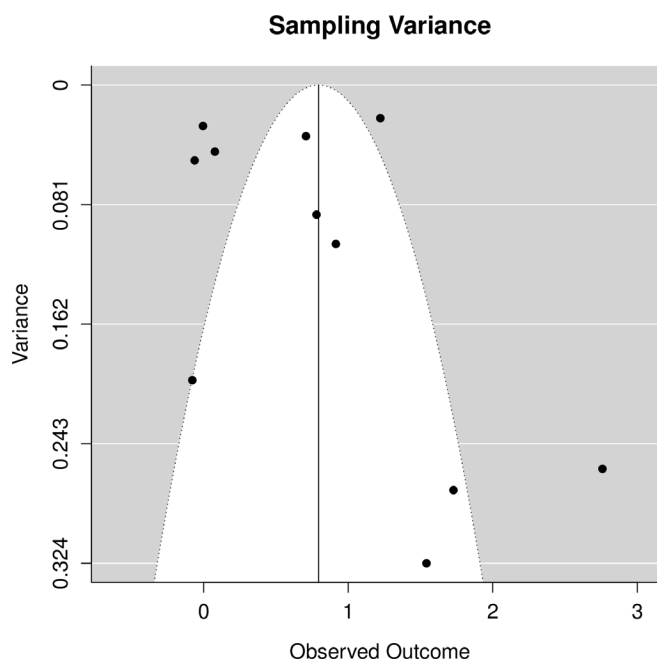


Figure 5 Funnel plot depicting risk of bias across studies included in the quantitative analysis ($k=10$).

this determination. The open areas of the plot may be due to sample size and might not be indicative of bias. Regardless, an equal number of points was found on both sides of the vertex, which provides evidence in favour of an absence of publication bias.

DISCUSSION

Summary of main findings

This is the first systematic review of observational and experimental designs on MT development and training in sport. Twelve studies were identified for inclusion in the review, only three of which did not meet criteria for high risk of bias. The accompanying meta-analysis that was performed with 10 studies revealed a large effect for interventions targeting MT development.

Although there has been a widely held belief that MT may be developed through intervention and training initiatives (see¹⁵), a mixture of positive (eg,³⁷) and null effects (eg,⁴¹) for MT training efforts have been reported in the literature. For the first time, the current review offers a synthesis of existing intervention research that supports the developmental potential of MT via specialised training approaches and programmes. This is an encouraging finding for sport personnel (eg, coaches, practitioners) who are interested in using theoretically grounded and evidenced-based approaches to safely and effectively promote MT among athletes they work with. However, wide variability in efficacy across studies and the high risk of bias identified for a majority of studies suggests that caution should be applied when deciding on the validity and practical utility of the MT training approaches reported in studies included in this review. While the potential to development MT is promising, there is little evidence to suggest that any particular approach to training MT in sport should be favoured over another. Thus, more rigorous research is needed to strengthen the body of evidence on the efficacy of MT development and training approaches before a determination can be made about which empirically based MT interventions ought to be prioritised when attempting to improve athletes' MT (see [figure 2](#)).

Implications for practice

There is currently insufficient evidence to advocate a specific approach to developing MT, and practitioners should be aware of the level of quality of MT research before adopting any of the interventions used in the recent literature. However, there are several findings of this review that may assist practitioners with avenues for cultivating MT in sport. First, two studies investigated the effect of a physical training-only approach on athletes' MT. Neither intervention was found to be effective in building MT. The remaining MT interventions were theoretically grounded in psychology (eg, psychological skills training), some of which were of moderate to high methodological quality and were found to be effective (eg,³⁵). Consequently, until a larger quantity of studies examining the effects of physical training approaches

(either alone or in combination with psychological approaches) on MT are amassed, it would be prudent to implement such approaches to develop MT with caution. Second, numerous qualitative studies have found that the natural developmental trajectory of MT tends to occur incrementally over a long-term period (see¹⁰). Only two studies included in this review employed longitudinal designs exceeding 6 months, but both of those studies reported large changes in MT. There was much greater variability in outcomes for the MT interventions that were briefer in duration. Thus, the likelihood of successfully improving may be enhanced if MT training initiatives are implemented gradually over a longer period of time (see [figure 2](#) for a summary of methodological limitations of existing research on MT interventions in sport and recommendations for future research).

Implications for research

The largest proportion of studies included in this review ($k=7$) used controlled designs to assess the efficacy of MT training approaches, but only three could be categorised as RCTs. There is a need to improve the methodological rigour of MT intervention research to address common sources of potential bias identified in this review, including the prioritisation of RCTs as a gold standard deductive approach for assessing the treatment efficacy of MT interventions.

Although psychological interventions impose some challenges to eliminating all risks of bias (eg, blinding of participants), random assignment, concealment of allocation, blinding of therapists and assessors of outcomes, and implementing measures to reduce attrition are critical areas to address in future controlled designs. Because multi-group experimental designs are not always feasible in competitive sporting environments, alternative methodologies (ie, single-group and single-subject designs) were frequently employed to evaluate the efficacy of MT training initiatives ($k=5$). Many of the factors that affected the quality of the CISs also applied to the alternative approaches used in studies that were reviewed (eg, non-blinded assessors of outcomes, low follow-up response rates). Multiple baseline and post-treatment assessments were also missing from all such studies. These factors ought to be strengthened in subsequent studies to maximise the utility of non-controlled trials and case study approaches for investigating MT interventions. Regardless of the methodology that researchers employ, improved reporting of methods is needed to enhance the interpretability and replicability of MT intervention research (see [figure 2](#)).

Considering the long-term developmental process of MT, it is surprising that longer post-intervention follow-up periods were not reported in any of the studies included in this review. To assess the long-term stability of improvements in MT acquired via training approaches, longitudinal post-treatment follow-ups should accompany future MT interventions. Collectively evaluating the sample characteristics of the studies reviewed,

it is apparent that female athletes have been under-represented in MT intervention research. Consistent with the general trend towards disproportionate sampling of male athletes that occurs in applied sport psychology research,⁴⁵ further research involving female athletes is needed to improve the generalisability of MT interventions beyond male athletes.

Although the studies reviewed included athletes participating in various sports, evidence on sport-specific differences in MT attributes suggests that interventions ought to be developed with consideration to the cultural and contextual nuances of MT that exist within each sport.^{3 46} Most studies included in this review ($k=8$) involved groups of athletes from more developed, higher-income countries that are characterised principally by individualistic value systems (ie, Australia, Canada, UK, USA). The efficacy of MT interventions has largely been unexplored among athletes from countries in Africa, Asia and South America. Prevailing value systems (eg, collectivistic principles) in many countries from those regions are likely to influence local understandings of MT¹ in ways that could affect how MT training in sport is approached (ie, cultural invariance of treatment approaches). As researchers continue to explore appropriate strategies to develop MT in athletes, it would be important to examine whether customised MT interventions that are sensitive cultural (eg, societal value systems, ideals of MT in specific sports) and contextual dynamics yield superior outcomes compared with more generic approaches to developing MT. Consistent with the findings of other reviews on MT (eg,¹¹¹), many studies included in this review used a single informant approach to measure MT ($k=10$) acquired via athlete self-report ($k=9$). With the exception of one study (ie,²³), MT was measured using instruments designed specifically to capture the psychological attributes of MT. Only three studies evaluated sport performance metrics alongside MT. Multi-informant approaches that assess psychological and behavioural indicators of MT could be integrated into future research to triangulate and more comprehensively evaluate the efficacy of MT interventions. Furthermore, studies that measure objective indicators of athletic performance are likely to offer supplementary evidence for adjudicating the internal validity of interventions that target MT^{1 4} (see figure 2).

Limitations

Despite the robust research design, there are also some limitations of this systematic review and meta-analysis. A thorough attempt was made to identify all relevant studies, but there is the possibility that relevant research (eg, unpublished technical reports, master's theses) was not discovered. Second, the number of studies that met the pre-established criteria for inclusion in the meta-analysis was 10, which limited the possible analytic techniques that could be used with adequate statistical power. For example, a higher quantity of studies would have enabled additional analyses of potential moderating variables (eg, length of training) to be performed,

which may have yielded insights into differences in the effectiveness of MT training approaches that were found across the studies reviewed. Third, MT was measured using psychological instrumentation. Measurement error may have influenced the results of the meta-analysis, but the expectation of bias was small because the instruments were associated with evidence of acceptable reliability.

Conclusion

In 2008, Connaughton and colleagues concluded that there was a need for researchers to abandon anecdotally based investigations and focus entirely on empirically based MT research.⁴⁷ This review is the first to synthesise existing empirical research on the effectiveness of MT training from observational designs and experimental designs. Therefore, the authors are unable to compare the results of this meta-analysis with others. Although a positive effect was observed for increasing levels of MT among a variety of sports-related groups as a result of one or more types of intervention, the quality of the studies reviewed should be considered alongside interpretations that are made about the results. Based on the evidence presented in this review, we extend the call made by Jones *et al*⁴⁸ approximately 20 years ago for a more scientific approach to the study of MT by emphasising the need to strengthen the methodological rigour in designing, implementing and evaluating the effectiveness of MT training approaches.

Author affiliations

¹Department of Sport and Wellness, SUNY Plattsburgh, Plattsburgh, New York, USA

²Department of Health, Exercise Science, and Recreation Management, University of Mississippi, University, Mississippi, USA

³Department of Educational Psychology, Baylor University, Waco, Texas, USA

⁴Department of Psychology, University of the Free State, Bloemfontein, Free State, South Africa

⁵Department of Nutrition, Food, and Exercise Sciences, Florida State University, Tallahassee, Florida, USA

Funding The authors have not declared a specific grant for this research from any funding agency in the public, commercial or not-for-profit sectors.

Competing interests None declared.

Patient consent for publication Not required.

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement All data relevant to the study are included in the article or uploaded as supplementary information.

Open access This is an open access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited, appropriate credit is given, any changes made indicated, and the use is non-commercial. See: <http://creativecommons.org/licenses/by-nc/4.0/>.

Author note References marked with an asterisk indicate studies included in the meta-analysis.

ORCID iD

Andreas Stamatis <http://orcid.org/0000-0001-9438-0558>

REFERENCES

- 1 Cowden RG. Mental toughness and success in sport: a review and prospect. *Open Sports Sci J* 2017;10:1–14.

- 2 Gucciardi DF, Peeling P, Ducker KJ, *et al.* When the going gets tough: mental toughness and its relationship with behavioural perseverance. *J Sci Med Sport* 2016;19:81–6.
- 3 Cowden RG, Crust L, Tibbert SJ, *et al.* Mental toughness development and training in sport. In: Bertollo M, Filho E, Terry PC, *et al.*, eds. *Innovations in mental skills training*. New York, NY: Routledge, 2020.
- 4 Gucciardi DF. Mental toughness: progress and prospects. *Curr Opin Psychol* 2017;16:17–23.
- 5 Coulter TJ, Mallett CJ, Singer JA. A three-domain personality analysis of a mentally tough athlete. *Eur J Pers* 2018;32:6–29.
- 6 Gucciardi DF, Hanton S, Gordon S, *et al.* The concept of mental toughness: tests of dimensionality, nomological network, and traitness. *J Pers* 2015;83:26–44.
- 7 Hardy L, Bell J, Beattie S. A neuropsychological model of mentally tough behavior. *J Pers* 2014;82:69–81.
- 8 Weinberg R, Butt J, Mellano K, *et al.* The stability of mental toughness across situations: taking a social-cognitive approach. *Int J Sport Psychol* 2017;48:280–302.
- 9 Gucciardi DF, Hanton S. Critical reflections and future considerations. In: Schinke RJ, McGannon KR, Smith B, eds. *Routledge international handbook of sport psychology*. New York, NY: Routledge, 2016: 439–57.
- 10 Anthony DR, Gucciardi DF, Gordon S. A meta-study of qualitative research on mental toughness development. *Int Rev Sport Exerc Psychol* 2016;9:160–90.
- 11 Lin Y, Mutz J, Clough PJ, *et al.* Mental toughness and individual differences in learning, educational and work performance, psychological well-being, and personality: a systematic review. *Front Psychol* 2017;8:1345.
- 12 Crust L, Clough PJ. Developing mental toughness: from research to practice. *J Sport Psychol Action* 2011;2:21–32.
- 13 Roebuck GS, Fitzgerald PB, Urquhart DM, *et al.* The psychology of ultra-marathon runners: a systematic review. *Psychol Sport Exerc* 2018;37:43–58.
- 14 Gucciardi DF, Gordon S, Dimmock JA, *et al.* Understanding the coach's role in the development of mental toughness: perspectives of elite Australian football coaches. *J Sports Sci* 2009;27:1483–96.
- 15 Weinberg R, Freysinger V, Mellano K. How can coaches build mental toughness? Views from sport psychologists. *J Sport Psychol Action* 2018;9:1–10.
- 16 Newport K. Maryland Strength Coach Rick Court Resigns in Wake of Jordan McNair's Death, 2018. Available: <https://bleacherreport.com/articles/2790984-maryland-strength-coach-rick-court-resigns-in-wake-of-jordan-mcnairs-death>
- 17 Wojtys EM. Houston mystery. *Sports Health* 2019;11:207–8.
- 18 Moher D, Liberati A, Tetzlaff J, *et al.* Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *Ann Intern Med* 2009;151:264–9.
- 19 Akobeng AK. Understanding systematic reviews and meta-analysis. *Arch Dis Child* 2005;90:845–8.
- 20 Hammersley M. On 'systematic' reviews of research literatures: a 'narrative' response to Evans & Benefield. *Br Educ Res J* 2001;27:543–54.
- 21 Liberati A, Altman DG, Tetzlaff J, *et al.* The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: explanation and elaboration. *PLoS Med* 2009;6:e1000100.
- 22 Higgins JPT, Green S. *Cochrane handbook for systematic reviews of interventions*. Chichester, UK: John Wiley & Sons Ltd, 2008.
- 23 Parkes JF, Mallett CJ. Developing mental toughness: attributional style retraining in rugby. *Sport Psychol* 2011;25:269–87.
- 24 Edwards P, Clarke M, DiGiuseppe C, *et al.* Identification of randomized controlled trials in systematic reviews: accuracy and reliability of screening records. *Stat Med* 2002;21:1635–40.
- 25 Rasti MA, Abubakar ZB, Abidin ZB, *et al.* The effect of mental training program on improving mental toughness among Tehran high school student athletes. *Asian Soc Sci* 2015;11:175.
- 26 Maher CG, Sherrington C, Herbert RD, *et al.* Reliability of the PEDro scale for rating quality of randomized controlled trials. *Phys Ther* 2003;83:713–21.
- 27 The University of Sydney. PEDro statistics, 2019. Available: <https://www.pedro.org.au/english/downloads/pedro-statistics/>
- 28 National Institutes of Health. Quality assessment tool for before–after (pre–post) studies with no control group. Available: <https://www.nhlbi.nih.gov/health-pro/guidelines/in-develop/cardiopulmonary-risk-reduction/tools/before-after>
- 29 Cook B, Buysse V, Klingner J, *et al.* Council for exceptional children: standards for evidence-based practices in special education. *Teach Except Child* 2014;46:206–12.
- 30 Cohen J. *Statistical power analysis for the behavioral sciences*. Hillsdale, NJ: Lawrence Erlbaum, 1988.
- 31 Hedges LV, Pustejovsky JE, Shadish WR. A standardized mean difference effect size for single case designs. *Res Synth Methods* 2012;3:224–39.
- 32 Viechtbauer W. Conducting meta-analyses in R with the metafor package. *J Stat Softw* 2010;36:1–48.
- 33 R Core Team. *R: a language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing, 2019.
- *34 Abdelbaky FM. Impacts of mental toughness program on 20 km race walking. *Sci Mov and Health* 2012;12:67–71.
- *35 Ajilchi B, Amini HR, Ardakani ZP, *et al.* Applying mindfulness training to enhance the mental toughness and emotional intelligence of amateur basketball players. *Australas Psychiatry* 2019;27:291–6.
- *36 Bhabri E, Dhillon PK, Sahni SP. Effect of psychological interventions in enhancing mental toughness dimensions of sports persons. *J Indian Acad Appl Psychol* 2005;31:65–70.
- *37 Bell JJ, Hardy L, Beattie S. Enhancing mental toughness and performance under pressure in elite young cricketers: a 2-year longitudinal intervention. *Sport Exerc Perform Psychol* 2013;2:281–97.
- *38 Castello M, Reed JP, Lund R, *et al.* Relationship between physical training, ratings of perceived exertion, and mental toughness in female NCAA division I volleyball players. *Sport J* 2018;21.
- *39 Drees MJ, Mack MG. An examination of mental toughness over the course of a competitive season. *J Sport Behav* 2012;35:377–86.
- *40 Gucciardi DF, Gordon S, Dimmock JA. Evaluation of a mental toughness training program for youth-aged Australian footballers: I. A quantitative analysis. *J Appl Sport Psychol* 2009;21:307–23.
- *41 Mahoney JW, Ntoumanis N, Gucciardi DF, *et al.* Implementing an autonomy-supportive intervention to develop mental toughness in adolescent rowers. *J Appl Sport Psychol* 2016;28:199–215.
- *42 Sheard M, Golby J. Effect of a psychological skills training program on swimming performance and positive psychological development. *Int J Sport Exerc Psychol* 2006;4:149–69.
- *43 Slack LA, Maynard IW, Butt J, *et al.* An evaluation of a mental toughness education and training program for early-career English football league referees. *Sport Psychol* 2015;29:237–57.
- 44 Gucciardi DF, Gordon S, Dimmock JA. Evaluation of a mental toughness training program for youth-aged Australian footballers: II. A qualitative analysis. *J Appl Sport Psychol* 2009;21:324–39.
- 45 Taylor J. *Assessment in applied sport psychology*. Champaign, IL: Human Kinetics, 2017.
- 46 Eubank MR, Nesti MS, Littlewood MA. A culturally informed approach to mental toughness development in high performance sport. *Int J Sport Psychol* 2017;48:206–22.
- 47 Connaughton D, Wadey R, Hanton S, *et al.* The development and maintenance of mental toughness: perceptions of elite performers. *J Sports Sci* 2008;26:83–95.
- 48 Jones G, Hanton S, Connaughton D. What is this thing called mental toughness? An investigation of elite sport performers. *J Appl Sport Psychol* 2002;14:205–18.