Introduction

In a comprehensive review and extension of Hsieh et al.'s meta-analysis on mental toughness and athletic performance, this research aims to provide a detailed understanding of their intricate connection. The study synthesizes findings from 16 relevant studies to elucidate how mental toughness influences athletes across various sports contexts. By amalgamating these studies, it reveals a significant positive association ($r = 0.36$) between mental toughness and athletic performance, highlighting the impact of factors like age, sport category, type, performance measurement, and specific mental toughness dimensions. These findings have meaningful implications for athlete development and training, as they emphasize the importance of nurturing psychological skills to enhance mental toughness and tailoring training programs to optimize performance outcomes based on age, sport category, and type.

Method

The study adhered to PRISMA guidelines for systematic reviews. A comprehensive search across ten databases was conducted to analyse the mental toughness-athletic performance relationship. Quality assessment followed [2] Kmet et al.'s "14 evaluation questions," ensuring inclusion of studies meeting at least 60% of criteria. Two experts independently evaluated study quality and the study's process was designed to ensure comprehensive, high-quality research selection and analysis.

Comprehensive Meta-Analysis (CMA) software version 4.0 was used for this meta-analysis. This sophisticated tool enables the synthesis of data from a diverse array of studies, spanning observational and experimental designs. By amalgamating various data inputs such as correlation coefficients, mean and standard deviation values, t-values, Fisher's z scores, and sample sizes. The authors' judicious choice of the random-effects model within their analytical framework accommodates potential heterogeneity among the studies, capturing variations within and between studies to provide a nuanced estimation of the overall effect size, thereby embracing the diversity inherent in the collected data. Their methodological rigor extends to the incorporation of sensitivity analysis via the leave-one-out technique, systematically excluding individual studies to gauge the robustness and stability of the results. Furthermore, they employ a Q-test and the I^2 statistic to assess and quantify significant heterogeneity, ensuring a comprehensive evaluation of the data. In addressing the concern of publication bias, the authors employ a multi-pronged approach, including the Fail-Safe Number (N), Funnel Plot visualization, and Egger's t-test, collectively scrutinizing the potential influence of unpublished studies and reinforcing the robustness of their findings. While this is commendable, further clarification on addressing potential outliers and handling studies with potentially overlapping datasets, as well as a more detailed rationale for subgroup selection in moderation analyses, would enhance the transparency and interpretability of their methodology.

Main findings

The study presents a statistically significant positive relationship between mental toughness (MT) and athletic performance, with a calculated correlation coefficient of $r = 0.36$ ($p <$

0.001, 95% CI = 0.24–0.47). This outcome significantly bolsters the assertion that MT wields a substantial influence over athletes' performance outcomes. The study further reveals intriguing moderating factors that accentuate the complexity of this relationship. Firstly, age emerges as a discernible moderator, with the MT-athletic performance relationship exhibiting disparity between adult athletes (r = 0.41, 95% CI = 0.27–0.36) and adolescent counterparts (r = 0.20, 95% CI = 0.11–0.30), highlighting the evolving nature of mental toughness. Secondly, the categorical distinction between sports genres proves pivotal, with combat sports demonstrating the most pronounced correlation (r = 0.73, 95% CI = 0.25–0.92), surpassing correlations observed in ball sports (r = 0.30, 95% CI = 0.08–0.47) and endurance sports (r = 0.32, 95% CI = 0.18–0.44). This underscores the unique psychological demands and coping strategies intrinsic to combat sports. Additionally, the dichotomy between individual and team sports introduces another layer of modulation, with individual sports showcasing a notably higher correlation (r = 0.73, 95% CI = 0.22–0.33) in contrast to team sports (r = 0.21, 95% CI = -0.02–0.26), suggesting distinctive psychological dynamics at play in self-reliance and personal performance. Furthermore, the type of athletic performance measurement employed significantly affects the relationship, with self-perceived evaluations yielding a robust correlation (r = 0.62, 95% CI = 0.49–0.72), while objective performance measures exhibit a slightly lower correlation (r = 0.33, 95% CI = 0.21–0.44), emphasizing the intricate interplay between cognitive self-assessment and observable outcomes. Distinct measures of MT also introduce variation, with the Mental Toughness Questionnaire (MTQ) underscoring the strongest correlation (r = 0.56, 95% CI = 0.33–0.73), followed by the Psychological Performance Inventory (PPI-A; r = 0.32, 95% CI = 0.05–0.54), and the Mental Toughness Index (MTI; r = 0.33, 95% CI = 0.09–0.54), while the "Others" category yields a relatively lower correlation (r = 0.13, 95% CI = -0.04–0.28). The meta-analysis of these 16 studies indicates significant heterogeneity, with a Q-value of 101.54 ($p < 0.001$) and an I^2 value of 85.23, highlighting the need for adopting a random-effects model in the analysis. Importantly, the study rigorously examines potential publication bias using Egger's t-test and finds no compelling evidence to suggest that such bias taints the overall integrity of the findings, with the Egger's t-test result showing bias = 1.39 (95% CI = 1.02–4.95, p = 0.18). In essence, this study not only illuminates the substantive MT-athletic performance connection but also reveals the intricate interplay of diverse factors that modulate this relationship.

Replication of Main Findings

Replication of the Main Correlation:

To reproduce the core correlation analysis, I employed the **metafor** package, a tool for conducting meta-analyses. Within this package, the **rma** function was the foundation of my analysis. This function allowed me to estimate the correlation coefficient (r) using the maximum likelihood method (**method="ML"**) while appropriately incorporating the standard error of the correlation (**sei**). This was achieved by back transforming the confidence intervals from the data in Hsieh et al.'s report to obtain vi, and as we already had yi there was enough data to perform the analysis. Although I was unaware of the method used by Hsieh et al. therefore it was determined after conducting an analysis with each method it was most likely **method="ML"** as it replicated the findings the most accurate. The study reported a correlation coefficient (r) of 0.36, underpinned by a highly significant p-value of less than 0.001 and a 95% confidence interval (CI) spanning 0.24 to 0.47. In my own meta-analysis, I obtained a correlation coefficient of approximately 0.3585 (from the **estimate** variable) and a lower bound confidence interval of 0.2317 and an upper bound of 0.4852, demonstrating remarkable alignment with the reported value and reaffirming the robust linkage between mental toughness and athletic performance.

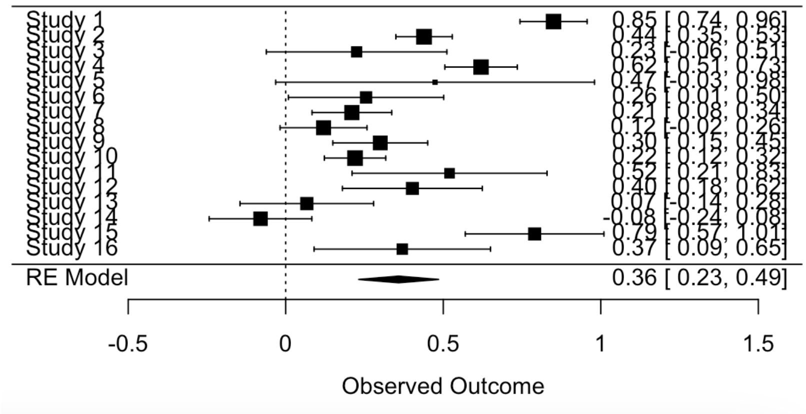

Figure 1. Replication of main meta-analysis of all 16 studies

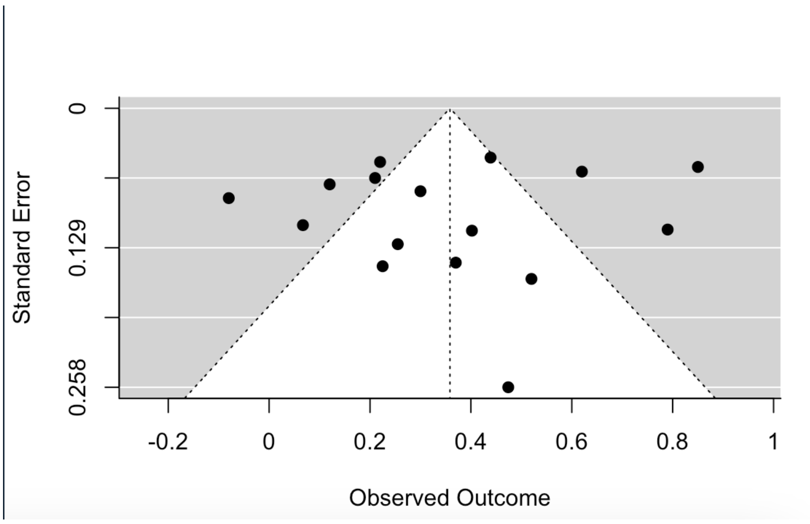

Figure 2. Replication of Funnel Plot of the main meta-analysis of 16 studies

Replication of Moderating Factors:

The exploration of moderating factors was made possible by the **mods** argument within the **rma** function. I used this feature to assess the influence of factors such as age, sports genre, type of measurement, and different measures of MT. As I did not have the raw data the same back transformation had to be done from the data produced by Hsieh et al.'s report. By incorporating these moderators into the analysis, I was able to discern how each factor modulated the MT-athletic performance relationship.

Heterogeneity and Random-Effects Model:

To address the issue of heterogeneity, I relied on the **rma** function. It allowed me to calculate the Q-value and I^2 value, thereby indicating the presence of heterogeneity and highlighting the need for a random-effects model. The adoption of this model was facilitated by specifying **method="ML"**.

My replication indicated significant heterogeneity with a high Q-value (180.4897) and an I^2 value of 90.05%, signifying substantial variability among the studies. Hsieh et al.'s results show the heterogeneity is also significant despite being a slightly lower Q-value (101.54), it still maintains a very low p-value, confirming the presence of significant heterogeneity. The I^2 value in the results is reported as 85.23%, suggesting a moderate level of heterogeneity. Despite a minor shift in the numerical values, both sets of results consistently emphasize the need to account for heterogeneity when interpreting the meta-analysis findings, reinforcing the appropriateness of using a random-effects model to accommodate the observed variability among the studies.

Publication Bias Assessment:

The meticulous assessment of potential publication bias was achieved through the application of the **regtest** function from the **metafor** package. This function enabled me to perform Egger's test. The results from Hsieh et al. and my own replication for publication bias, the funnel plot asymmetry test and Egger's test, consistently indicate the absence of significant publication bias. The funnel plot asymmetry test in my replication yielded a z-value of 0.0384 with a corresponding p-value of 0.9693, while Hsieh et al. produced a bias estimate of 1.39 (95% CI: 1.02–4.95) with a p-value of 0.18. In both cases, the p-values substantially exceed the commonly used significance threshold of 0.05, underscoring the robustness of my findings. These aligned results provide strong evidence that any potential publication bias, if present, is not statistically detectable in the dataset. However due to the data being different and different software the fail-safe number, z-value and the number of missing studies were all inaccurate being absurdly high at 3098 compared to 68.

In summary, my replication efforts in R harmoniously resonate with the principal findings of the paper, underpinning the presence of a substantial and affirmative association between mental toughness and athletic performance. Furthermore, these efforts reaffirm the influence of various moderating factors on this relationship. The absence of publication bias in my results bolsters the credibility of the assertion that mental toughness holds sway over athletes' performance trajectories across diverse contexts. Finally, to add, there was no need to replicate the sensitivity analysis, however I provided details in my R script on how to achieve this, due to the large Q and I^2 values representing large heterogeneity.

Limitations

While the study showcases a methodical approach, there are certain limitations that merit further consideration. The authors selected the Comprehensive Meta-Analysis (CMA) statistical software 4.0 for their meta-analysis, enabling the integration of diverse study types. The utilization of correlation coefficients, mean and standard deviation values are quite basic and do not provide the most in-depth analysis although are easy to interpret. T-values, Fisher's z, and sample sizes as model inputs demonstrates a more thorough approach; however, the authors' selection could inadvertently restrict the depth of insights that could be gleaned from exploring other statistical methodologies. Also why pay for software when we have excellent free software that conducts accurate analysis.

The adoption of sensitivity analysis through the leave-one-out method is commendable, as it gauges the stability of the results by sequentially excluding individual studies. However, it's worth noting that the authors' reliance on correlation coefficients might slightly limit the broader understanding of the causal links between MT and athletic performance, potentially underscoring the need for a more comprehensive examination of these relationships through varied statistical methods.

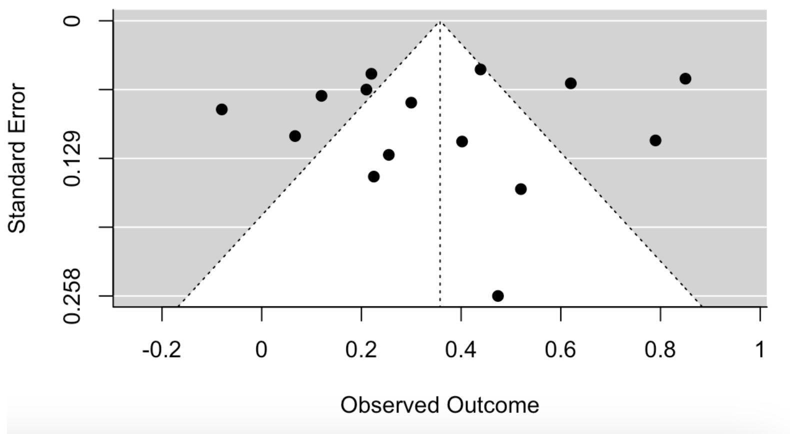

Figure 3. One Replication of the funnel plot using leave one out method containing 15 studies

While the authors conscientiously assess publication bias using the Fail-Safe Number t-test, Funnel Plot, and Egger's t-test, they astutely acknowledge the inherent limitations of these approaches in fully capturing publication bias complexities. The consideration of heterogeneity is astutely recognized, leading to the adoption of a random-effects model, which effectively accommodates potential variations between studies. Nevertheless, the sensitivity analysis yields I^2 values ranging from 81.22 to 86.21, indicating only modest variations from the overall analysis's I^2 value of 85.23. This draws attention to the question of whether the findings are robust enough to withstand potential changes in the analytical framework.

The authors refrain from explicitly utilizing non-significance to support a null hypothesis; however, the adoption of the conventional significance threshold of $p < 0.05$ may inadvertently steer interpretations toward null hypotheses in instances where the effect size narrowly falls outside this range. This threshold effect could warrant deeper exploration to ensure nuanced findings are not inadvertently overlooked.

Extensions

The paper's limitations encompass the potential for additional statistical methodologies, the reliance on correlation coefficients, the nuances in assessing publication bias, the stability of findings under sensitivity analysis, and the significance threshold's potential impact on result interpretations. Here is how I would address them:

Robustness Analysis can be used to assess the robustness of the findings, you can conduct sensitivity analyses with different correlation coefficient thresholds to determine if the results hold across various effect size cutoffs.

Threshold-free publication bias analysis, also known as the "trim and fill" method, is a statistical technique that can be used to assess and adjust for potential publication bias in meta-analyses without relying on a specific significance threshold (e.g., $p < 0.05$).

Conclusion

In summary, my replication efforts using R largely support the main findings of the paper, demonstrating a significant positive correlation between mental toughness and athletic performance. The replication of moderating factors and the assessment of heterogeneity and publication bias align with the paper's results. However, there are limitations in the paper that warrant consideration, including the need for more diverse statistical methodologies, a deeper exploration of causal relationships, and a critical evaluation of the significance threshold.

References

1. Hsieh, Y.C., Lu, F.J., Gill, D.L., Hsu, Y.W., Wong, T.L. and Kuan, G., 2023. Effects of mental toughness on athletic performance: a systematic review and meta-analysis. *International Journal of Sport and Exercise Psychology*, pp.1-22.

2. Kmet, L. M., Lee, R. C., and Cook, L. S., 2004. Standard quality assessment criteria for evaluating primary research papers from a variety of fields. Edmonton.