

Review article:

THE EFFECTS OF CHALLENGE AND THREAT STATES ON PERFORMANCE OUTCOMES: AN UPDATED REVIEW AND META-ANALYSIS OF RECENT FINDINGS

Adrian Hase^{a,*} , Maximilian Nietschke^b , Maciej Kłoskowski^c , Kacper Szymański^c ,
Lee Moore^d , Jeremy P. Jamieson^e , Maciej Behnke^c 

^a Medicine Section, University of Fribourg, Fribourg, Switzerland

^b Institute of Philosophy, University of Bern, Bern, Switzerland

^c Adam Mickiewicz University, Poznań, Poland

^d University of Bath, Bath, United Kingdom

^e University of Rochester, Rochester, USA

* **Corresponding author:** Adrian Hase, PhD, University of Fribourg, Chemin du Musée 8, 1700 Fribourg, Switzerland. E-mail: adrian.hase@unifr.ch

<https://dx.doi.org/10.17179/excli2024-7995>

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0/>).

ABSTRACT

The biopsychosocial model (BPSM) of challenge and threat provides a framework for understanding stress responses in motivated performance situations, including how stress relates to performance. In this model, experiences of challenge—characterized by evaluations of personal coping resources matching or exceeding situational demands—elicit approach-oriented patterns of physiological responding and tend to facilitate performance, whereas threat—characterized by demands exceeding resources—elicit avoidance-oriented patterns of physiological responding and tend to impair performance. Extant systematic reviews and meta-analyses support the idea that challenge facilitates performance relative to threat (Behnke and Kaczmarek, 2018; Hase et al., 2019c). The present systematic review and meta-analysis builds on this evidence base by examining whether conclusions replicate in recent research (post-2017), which is important given seismic cultural shifts tied to a worldwide pandemic, civil unrest, and skyrocketing mental health problems tied to stress. The analysis included 62 studies published between 2017 and 2024 (total $N = 7,418$ participants). The meta-analytic findings indicate that individuals in a challenge state achieve better performance outcomes than those in a threat state across multiple domains (e.g., education, sport). While effect sizes were small, the risk of bias was generally low. These results reaffirm the utility of the BPSM and emphasize the importance of stress responses in influencing performance outcomes. These data also have the potential to inform future research on this topic by shedding light on expectable effect sizes and highlighting potential influences of publication bias and replicability issues.

Keywords: Challenge-threat index, demand resource evaluations, motivated performance, biopsychosocial, cognitive appraisal, psychophysiology

THE EFFECTS OF CHALLENGE AND THREAT STATES ON PERFORMANCE OUTCOMES: AN UPDATED REVIEW AND META-ANALYSIS OF RECENT FINDINGS

Psychophysiological stress responses are crucial for determining performance out-

comes across myriad high-pressure contexts, including athletics, education, and healthcare domains, to name a few (Hase et al., 2019c). Understanding the processes that undergird individual psychophysiological stress responses in such contexts can help prevent underperformance and optimize outcomes (e.g.,

mental and physical health; McLoughlin et al., 2024). The Biopsychosocial Model (BPSM) of Challenge and Threat (CAT) (e.g., Blascovich and Tomaka, 1996) integrated concepts from Lazarus and Folkman's (1984) appraisal theory of stress as well as Dienstbier's (1989) theory of physiological toughness to offer a comprehensive framework of stress responses in motivated performance situations, or situations that are goal-relevant, evaluative, and potentially stressful. The BPSM has been applied across a range of fields, including sports (Moore et al., 2012, 2013a, b, 2015, 2018), education (Chalabaev et al., 2012; Jamieson et al., 2016; Schneider, 2004), and medicine (Moore et al., 2014; Vine et al., 2013). Here, we aimed to summarize the current findings on the role of CAT states in the performance context, updating our previous systematic review (Hase et al., 2019c) and a meta-analysis (Behnke and Kaczmarek, 2018).

Biopsychosocial model of challenge and threat

The BPSM presents the role of psychophysiological stress responses in motivated performance situations that require instrumental cognitive responses and often involve social evaluation (Blascovich and Mendes, 2010). A fundamental principle is the concept that cognitive evaluations of situational demands and personal coping resources operating in an integrated manner determine challenge- and threat-type stress responses in motivated performance contexts (Blascovich and Mendes, 2010; Jamieson, 2017; Mendes and Park, 2014). The BPSM conceptualizes “demands” and “resources” as multidimensional constructs.

Demand evaluations may include uncertainty, danger, and effort, which can be independent or intertwined. Similarly, resource evaluations encompass factors such as familiarity, knowledge, skills, dispositional traits, and social support, which may also overlap. Importantly, resource and demand evaluations can fluctuate independently (i.e., be ontologically distinct) or function as bipolar

factors affecting both processes (Seery, 2011). For instance, familiarity and uncertainty evaluations can simultaneously influence perceptions of both resources and demands (Blascovich, 2008; Seery, 2011). These evaluations determine whether an individual experiences a challenge- or threat-type response in a performance situation. A challenge response occurs when coping resources are appraised as matching or exceeding situational demands, whereas a threat response emerges when demands are perceived to outweigh coping resources.

While CAT states are often described as distinct, they are better understood as endpoints on a continuum of potential responses rather than as binary states (Seery, 2011). According to the BPSM, patterns of CAT cognitive evaluations determine physiological responses to stressors in performance contexts (see Mendes and Park, 2014, for a review). These stress responses are derived from activation of the sympathetic-adrenal-medullary (SAM) and hypothalamic–pituitary–adrenal (HPA) axes. Both CAT states are accompanied by SAM activation, leading to the synthesis and secretion of catecholamines like epinephrine and norepinephrine, which increase ventricular contractility (i.e., speeding up heart rate) and dilate blood vessels (Brownley et al., 2000). Both challenge- and threat-type stress responses are characterized by high sympathetic arousal. Thus, SAM-dominated challenge-type physiological responses are characterized by increased cardiac output – a volumetric measure of blood pumped by the heart across time – and decreased resistance in the peripheral vasculature (termed total peripheral resistance; Blascovich, 2008). These changes allow for the effective delivery of oxygenated blood flow to the brain and peripheral sites (e.g., muscles). Challenge-type physiological responses also allow for a rapid onset and offset of stress responses: resources are mobilized rapidly for performance, and individuals return to homeostasis quickly after performance is completed (Seery, 2011). Threat-type responses, however, not only elicit SAM

activation but also activate the HPA axis, which produces a more prolonged stress response due to the longer half-life of cortisol (the hormonal product of HPA activation). HPA activation tempers SAM effects and results in reduced (or little change in) cardiac output and increased resistance in the peripheral vasculature to concentrate blood in the core of the body in preparation for anticipated damage or defeat (Blascovich and Mendes, 2010; Jamieson, 2017).

More downstream, CAT responses have important consequences for decisions, behavior, and performance (Blascovich, 2008; Seery, 2013). For instance, CAT states exhibit differential patterns of motivational orientation, much like other affective states (e.g., anxiety and avoidance or anger and approach; Blascovich, 2008). In performance situations, a challenge state is generally associated with approach motivation and a threat state with avoidance motivation (e.g., Behnke et al., 2022; Jamieson and Mendes, 2016; Jamieson et al., 2014, 2022). Moreover, whereas a challenge state is typically associated with positive behavioral and performance outcomes (see Behnke and Kaczmarek, 2018; Hase et al., 2019c, for reviews), a threat state often impairs decision-making in the short-term and in the long-term is associated with accelerated “brain aging,” cognitive decline, poorer mental health (e.g., more symptoms of anxiety and depression) and cardiovascular disease (Jefferson et al., 2010; McLoughlin et al., 2024; Matthews et al., 1997; Yeager et al., 2022).

BPSM applications

The application of the BPSM to different research settings has led to support for its conceptualization of psychophysiological stress responses in motivated performance settings as well as the predictive utility of CAT responses for understanding performance outcomes. Precisely, a systematic review (Hase et al., 2019c) and a meta-analysis (Behnke and Kaczmarek, 2018) both found a challenge state to generally result in better performance than a threat state. The reviewed studies took

place in diverse performance contexts, including sport, medicine, academia, and aviation (e.g., Blascovich et al., 2004; Moore et al., 2013b, 2014; Seery et al., 2010; Vine et al., 2015). They relied on various research designs and included both CAT cognitive evaluations and cardiovascular responses. However, there were some inconsistencies in the reviewed research (e.g., the inclusion of studies until 2017 or 2018 and the differences between included studies), and the only meta-analysis to date restricted its focus to CAT cardiovascular responses and did not synthesize results relating to cognitive CAT evaluations and performance outcomes (Behnke and Kaczmarek, 2018). Moreover, societal changes tied to the COVID-19 pandemic and its aftermath warrant a new assessment of how psychophysiological stress responses map onto performance. Notably, stress processes related to anxiety, burnout, and political polarization have contributed to the growing “mental health pandemic” in the wake of COVID-19 (Daniali et al., 2023; Sacco, 2023). Thus, we argue that a new meta-analysis of both CAT cognitive evaluations and cardiovascular responses is much needed, especially given the numerous new studies that have been published since the last review by Hase and colleagues published in 2019c (e.g., Behnke et al., 2022, 2024; Jewiss et al., 2023, 2024; Sharpe et al., 2024a).

These new studies used established and novel approaches to examine CAT states from a BPSM perspective. For example, one study used an observational study to extend CAT research into the esports context, finding that CAT cognitive evaluations predicted goal ratio in the soccer game FIFA 19 (Behnke et al., 2022). Another study involved naturalistic observation of cardiovascular CAT responses in a pressurized performance context (i.e., competitive football or soccer match), which were then used to divide an elite-level sample into three cardiovascular reactivity groups: challenge, threat, and blunted reactivity (Dixon et al., 2020). The blunted reactivity group underperformed compared to the other two groups, which highlights the importance

of examining boundary conditions in BPSM-based CAT research (Hase et al., 2020). Finally, replication studies have been conducted, with Jewiss et al. (2023) finding similar benefits (i.e., of a challenge over a threat state) on cricket batting performance and Jewiss et al. (2024) failing to replicate the seminal findings of Blascovich et al. (2004) when connecting CAT states with season-long cricket batting performance.

Beyond these observational studies, several experiments have attempted to manipulate CAT states directly or indirectly and thereby improve performance. Indeed, in one study (McGreary et al., 2020), elite-level junior cricket batsmen were manipulated into a challenge or a threat state using established CAT instructions (from Moore et al., 2012). In the challenge group, all participants completed a subsequent cricket task successfully, whereas only one participant in the threat group completed the task successfully (McGreary et al., 2020). Another experiment tested a nutritional supplement-based manipulation of CAT states using the catecholamine precursor l-tyrosine (Hase et al., 2019a). Although acute l-tyrosine intake did not affect CAT states directly, it benefited individuals with relatively more threatened cardiovascular responses in maintaining optimal performance. Another study, which was a preregistered report, introduced a synergistic mindsets intervention inspired by past work (Yeager et al., 2022) to change esports players' CAT cognitive evaluations and cardiovascular responses (Behnke et al., 2024). After two weeks of training and applying the intervention, the authors did not find a relationship between CAT states and performance, which is notable due to the methodological strength of this study (i.e., among others, a relatively large sample size and preregistration).

In addition to the ongoing research and application of the BPSM in various performance contexts (e.g., education; Jamieson et al., 2021), the theoretical interest in understanding CAT states has spurred the development of new models that build on the BPSM. For instance, Meijen et al. (2020) proposed a

revised conceptual model that built on the theory of challenge and threat states (TCTSA) by Jones et al. (2009) and addressed CAT states in athletes (e.g., proposing new physiological markers such as neuropeptide Y and oxytocin). Similarly, Uphill et al. (2019) critically reviewed existing literature and suggested an alternative framework for understanding CAT states (e.g., that performers can be challenged, threatened, both, or neither in motivated performance situations). Finally, Hase et al. (2020) proposed adding blunted cardiovascular reactivity to the model to represent deficient task engagement and, more generally, motivational disengagement due to threat states. While these new models offer valuable theoretical advancements, they do not directly provide additional insights into the specific cognitive and cardiovascular pathways examined in the present review and are, therefore, not considered in detail for this meta-analysis.

Given the continued research interest in the BPSM to explain performance outcomes (e.g., Kelley et al., 2019), this meta-analytic review synthesizes the research published since 2017 to update the results following an uncertain time for society (e.g., due to COVID-19) and to build upon and scrutinize the conclusions of the previous reviews (i.e., Behnke and Kaczmarek, 2018; Hase et al., 2019c). Specifically, it aims to clarify the relationship between different CAT cognitive evaluations and cardiovascular responses and various performance outcomes across different research designs (e.g., experimental to observational). The focal hypothesis is that a challenge state would be associated with better performance outcomes than a threat state. We also performed exploratory analyses to test for potential moderators of the hypothesized effect, as well as examine associations among different measures of CAT (cognitive and physiological) to better understand the coherence among the different measure modalities. In doing so, this novel review provided information about how methodological choices (e.g., measuring cardiovascular responses before vs. during performance or

manipulating vs. observing CAT states) affect the CAT-performance relationship.

METHOD

Study selection

We used the PRISMA2020 application to graphically summarize the study selection process in a PRISMA flowchart (Haddaway et al., 2022; Figure 1). After the initial literature search (conducted in March 2024), the primary study selection process included three steps, which followed the procedure of Hase et al. (2019c), that is: screening based on title, followed by abstract, and finally full text. Three independent assessors (AH, MN, and MB) conducted the literature search, compared their results, and resolved any discrepancies by discussion (initial agreement in steps 1, 2, and 3 = 50.0 %, 56.8 %, and 81.7 %, respectively). The included databases were: MedLine, PsycINFO (combined in one search), SPORTDiscus, and Web of Science (in separate searches). The search terms were (“challenge and threat” AND “performance”). We then scanned the reference lists of the included articles for other articles that were potentially not picked up in the systematic search process. Furthermore, we included three studies (Moore et al., 2018; Sammy et

al., 2017; Scheepers, 2017) from a previous systematic review (Hase et al., 2019c), as the previous meta-analysis (Behnke and Kaczmarek, 2018) did not include these studies. Finally, we conducted a forward search to evaluate studies that cited those already identified as relevant to this project, by using the “cited by” option in Google Scholar. We excluded studies that: (1) did not measure performance outcomes objectively; (2) did not measure CAT cognitive evaluations (i.e., demands and resources) and/or cardiovascular indexes (i.e., CO, TPR, CTI); and (3) did not provide the effect sizes for the associations between CAT states and performance outcomes. We did not pre-register this meta-analysis.

Risk of bias analysis

To assess methodological quality, we used the Risk of Bias Assessment Tool for Nonrandomized Studies (RoBANS 2; Kim et al., 2013) and the revised Cochrane tool for assessing risk of bias in randomized trials (Sterne et al., 2019). We used the RoBANS 2 tool for the risk of bias assessment in observational studies that provided tests of associations between CAT states (cognitive evaluations and/or cardiovascular responses) and performance outcomes. The RoBANS 2 tool provides a domain-specific assessment of six

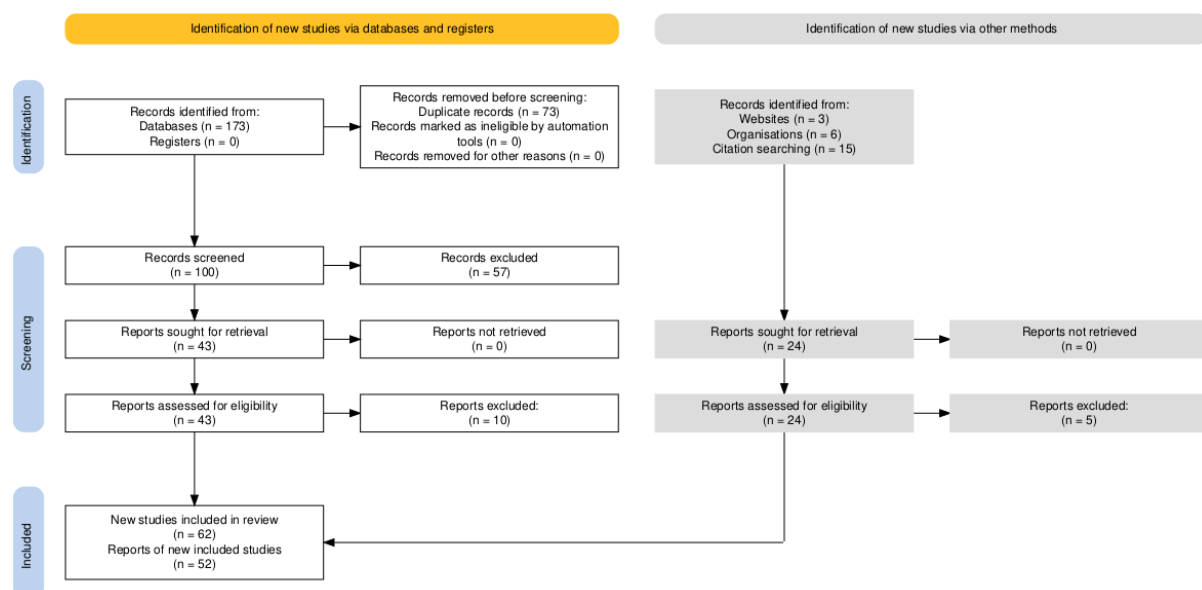


Figure 1: PRISMA flowchart

risk of bias domains: participant selection, confounding variables, measurement of exposure, blinding of outcome assessment, incomplete outcome data, and selective reporting. Possible ratings are low, unclear, or high risk of bias for each domain and the overall study assessment.

The only exceptions were studies that randomized participants into challenge versus threat groups, intentionally manipulated CAT states (e.g., with a set of instructions or nutritional supplement), and subsequently confirmed the successful manipulation of CAT states with a check using established cognitive or physiological CAT measures (e.g., Behnke et al., 2020). For these randomized studies, we used the revised Cochrane tool for assessing the risk of bias (ROB-2; Sterne et al., 2019). The ROB-2 comprises evaluations of bias in five distinct domains: bias due to issues with the randomization process, deviations from intended interventions, missing outcome data, outcome measurement, and selection of the reported result. Possible ratings are “low risk of bias”, “some concerns”, and “high risk of bias”. In both the ROBANS 2 and the ROB-2 tool, the highest individual domain rating also produced the overall risk of bias rating. To visualize the results of the risk of bias assessments, we created traffic light plots using *robvis* R package (McGuinness and Higgins, 2021).

Coding. Based on our previous reviews of the CAT literature (Behnke and Kaczmarek, 2018; Hase et al., 2019c), the following potential moderators were coded: (1) the timing of cardiovascular measures (before/during performance); (2) age of the sampled participants (sample mean in years); (3) gender of the sampled participants (percentage of females), (4) publication year, and (5) risk of bias.

When coding for the correlations with performance outcomes, we multiplied the effects by -1 for the types of tasks where lower numbers indicated better performance (e.g., time, error); thus, for all correlations, the higher the positive correlation coefficient, the more strongly the CAT cognitive evaluations

and cardiovascular responses correlated with better performance. Similarly, for the CAT cognitive evaluations, we multiplied the effects related to evaluated demands by -1; thus, for all CAT cognitive evaluations, the higher the positive correlation coefficient, the more strongly the challenge evaluations (i.e., personal coping resources match or exceed situational demands) correlated with better performance.

Effect size extraction. For the meta-analysis, we used the correlation coefficients between CAT cognitive evaluations and/or cardiovascular responses and performance outcomes as a measure of the effect size. For most studies, the authors reported the correlation coefficients. For studies reporting other metrics (e.g., between-group differences), we sent requests to authors to provide the correlation coefficients. Of the 47 authors we contacted, 37 responded to our inquiry (79 %), and of these, 29 sent us the requested data (78 %). No authors denied us access to the requested data. The list of authors who provided and did not provide the data is to be found in Supplementary data (“*Data.xlsx* file, “*mails*” sheet).

Meta-analytic procedures

We conducted meta-analyses using R (R Core Team, 2023) and the *metafor* package (Viechtbauer, 2010), adhering to established guidelines (Assink and Wibbelink, 2016; Viechtbauer, 2010). Given the anticipated high heterogeneity of effects (as in Behnke and Kaczmarek, 2018), we used a random-effects model. Random-effects models are recommended for meta-analyses because they facilitate the generalization of corrected effect sizes to the broader population (Field and Gillett, 2010; Schmidt and Hunter, 2015). Additionally, simulation studies have shown that using separate three-level models for different outcome types is the most reliable method for obtaining accurate standard error estimates (Fernández-Castilla et al., 2021). Following prior research (e.g., Behnke et al., 2022; 2023), we applied restricted maximum likelihood estimation to compute pooled mean

effect sizes. Several of the included studies reported multiple effect sizes. To address the dependency among these effect sizes, we used a three-level meta-analytic approach (Assink and Wibbelink, 2016; Viechtbauer, 2010). This method accounts for three distinct sources of variance: (1) between-study variance, which captures differences in effect sizes across studies (level three); (2) within-study variance, arising from multiple effect sizes reported in the same study (e.g., using two cardiovascular measures; level two); and (3) sampling variance of the extracted effect sizes (level one) (Cheung, 2014; Van den Noortgate et al., 2013).

Magnitude and consistency of the effects. We aimed to examine a pooled mean relationship between CAT cognitive evaluations and/or cardiovascular responses and performance outcomes. We interpret the results considering two parameters, namely, relationship magnitude (size of the mean effect sizes) and consistency (no/low heterogeneity vs. high heterogeneity). Heterogeneous effect sizes may indicate that: (a) the average relationship is not consistent for CAT states and performance outcomes; (b) the average relationship between CAT states and performance outcomes is moderated by different types of characteristics (e.g., timing); or (c) the size of the effect reflects real, contextual changes in the relationship between CAT states and performance outcomes.

We assessed the homogeneity of the calculated mean effect sizes using I^2 -statistic. A rejection of the null hypothesis suggests the presence of methodological or population differences that contribute to variability between studies. Additionally, the I^2 -statistic facilitates comparisons across meta-analyses (Higgins et al., 2003). It measures the proportion of variance attributable to within-study effects (level two of the meta-analytic model) and between-study effects (level three). The total I^2 value represents the overall variability, where 0 % indicates no dispersion and higher values signify varying degrees of heterogeneity: 0–40 % might be negligible, 30–60 % may indicate moderate heterogeneity, 50–

90 % may reflect substantial heterogeneity, and 75–100 % might represent considerable heterogeneity (Higgins et al., 2019). To test the significance of variance components at levels two and three, we conducted separate one-tailed log-likelihood ratio tests. These compared the deviance of the full model to reduced models that excluded one variance parameter. We then calculated the prediction intervals for the mean correlations to address the heterogeneity of the pooled effect sizes (Borenstein et al., 2017; Higgins et al., 2009; Riley et al., 2011). The prediction intervals present where the true effect size in 95 % of all comparable populations would fall. We intended to present the pooled effect sizes of the relationship between CAT states (cognitive evaluations and/or cardiovascular responses) and performance outcomes, as well as the prediction intervals, as correlation coefficients.

Publication bias. We conducted five publication bias analyses to determine whether null or weak results might have been systematically underrepresented in the literature (Schmidt and Hunter, 2015). Investigating potential bias in synthesized effect sizes is crucial, as studies reporting non-significant or negative findings are less likely to appear in peer-reviewed journals (Borenstein et al., 2009). Our strategy included five complementary methods: (1) a visual inspection of the funnel plot, (2) a rank correlation test, (3) Egger's regression test (Egger et al., 1997), (4) the trim-and-fill analysis (Duval and Tweedie, 2000), and (5) a moderator test in which the publication year of studies was tested as a moderator of the overall effect. This approach aligns with similar strategies employed in previous research (e.g., Assink et al., 2018, 2019; Behnke et al., 2022, 2023). In our strategy, we used some standard methods for two-level meta-analysis because their counterparts for three-level meta-analyses have not yet been developed and tested (Fernández-Castilla et al., 2020).

First, we visually inspected a funnel plot in which effect sizes are plotted against their standard error around an estimated summary effect (Egger et al., 1997). In contrast to large

studies, studies using small samples tend to produce effect sizes of different magnitudes due to increased variability in their sampling errors. Thus, effect sizes from smaller studies are expected to scatter widely at the bottom of the funnel plot. In contrast, effect sizes from larger studies are expected to be more concentrated at the top of the plot. As null or weak results are likely to be systematically suppressed from publication, an asymmetric distribution of effect sizes may be expected in the sense that effect sizes may be missing at the (bottom) left of the estimated summary effect in the funnel plot. In contrast, a symmetric effect size distribution with effect sizes equally distributed to the left and right of a summary effect would suggest the absence of publication bias. Second and third, we applied an adapted version of Egger's regression test and Begg and Mazumdar's rank-order correlation test (Assink et al., 2018, 2019). In the adapted Egger's test, we regressed effect sizes on their standard errors within a three-level meta-analytic framework, accounting for the dependency among effect sizes derived from the same primary studies. Unlike the classical Egger's test, this adaptation addresses effect size dependency. In Egger's regression test, a significant slope suggests the presence of publication bias, while in Begg and Mazumdar's rank-order test, a significant rank association indicates bias (Assink et al., 2018, 2019).

Fourth, we used the trim-and-fill method (Duval and Tweedie, 2000) to examine bias-corrected effect sizes. This method addresses funnel plot asymmetry by imputing effect sizes from "missing" studies to restore symmetry in the plot. Fifth, we tested for a potential "decline effect" (Schooler, 2011) by examining whether effect sizes decreased over time. This was done by regressing the summary effect on the publication year of the primary studies. A tendency for effect sizes to diminish over time may reflect what Ioannidis (2005) described as the law of initial results. To extend the range of publication years in this analysis, we included data from previous

meta-analyses (Behnke and Kaczmarek, 2018).

Moderator analyses. Finally, to examine potential moderators that may influence the relationship between CAT states and performance outcomes, we ran five types of separate moderator analyses. We tested whether the findings differ for different groups of studies: for example, studies that measured CAT cardiovascular responses before vs. during performance and studies with low, unclear, or high risk of bias. Furthermore, we explored whether the age of the sampled participants influenced the size of the effect sizes (sample mean in years), gender of the sampled participants (percentage of females), and publication year. We examined whether the relationship between CAT states and performance outcomes varied across (categories of) the factor by interpreting the results of an omnibus test. A significant *F* value in this test indicates differences across (categories of) the factor, serving as a potential moderator. To control for Type I error due to multiple comparisons, such as testing five different moderators within the same dataset—a common issue in meta-analyses (Cafri et al., 2010)—we applied Bonferroni correction to adjust probability values (Abdi, 2007). Following our previous meta-analyses (Behnke et al., 2022; 2023), we considered the mean effect of a moderator category only when at least three studies supported it.

RESULTS

Descriptive analyses

We identified 72 records in the first search (PsycINFO and Medline), 23 records in the second search (Sportdiscus, 22 duplicates), and 78 in the third search (Web of Science, 51 duplicates). Thus, 100 unique titles were evaluated in the first review step. Based on the initial evaluation, we selected 95 titles for inclusion in step 1 (50.0 % initial agreement between the three reviewers), 81 abstracts in step 2 (56.8 % agreement), and 43 full texts in step 3 (81.7 % agreement). We added three additional articles based on a reference list search, three studies from 2017, six sets of

unpublished results or published articles sent by contacted authors, and 12 reports identified by forward search. Thus, the final number of articles coded for the meta-analysis was 67.

After excluding papers that did not present the data needed for meta-analysis, we included 52 articles with 62 studies (Table 1), presenting 336 effect sizes obtained from 7,418 participants (14,220 performances) with the mean ages ranging from 13.26 to 72.00 years ($M = 21.70$ years, $SD = 7.66$; Figure 1). Sample sizes ranged from 10 to 499, with a mean sample size of 119.65 participants ($SD = 111.34$). More than half of the participants (51 %) were male. Most samples contained both genders, but eight samples were all-male (Behnke et al., 2020, 2022, 2024; Dixon et al., 2020; Jewiss et al., 2023, 2024; Sharpe et al., 2024b; Trotman et al., 2018), and three samples were all-female (Petzel and Casad, 2022; Scheepers, 2017; Turner et al., 2021). The included studies were published between 2017 and 2024 (median publication year = 2022).

The predominantly studied type of performance was math ($n = 9$ studies; 15 %), followed by esports (8; 13 %), throwing (5; 8 %), academic exams (5; 8 %), darts (4; 6 %), problem-solving tasks (4; 6 %), soccer (3; 5 %), golf (3; 5 %), indoor cycling (2; 3 %), presentation or speech performance (2; 3 %), cricket (2; 3 %), memory task (2; 3 %), and communication task, cooperation computer game, medical test, muscular endurance, netball, pattern recognition task, psychomotor vigilance task, running, shooting, tennis, Tetris, trampoline, and visual search task (all 1; < 2 %). The most frequent cardiovascular measure was Challenge Threat Index (CTI) (73 effect sizes; 36 %), followed by cardiac output (CO) (68 effect sizes; 34 %), and total peripheral resistance (TPR) (61

effect sizes; 30 %). The most frequent CAT cognitive evaluation measure was different adaptations of the two-item cognitive appraisal ratio (Tomaka et al., 1993) ($n = 27$ studies; 60 %), followed by the four-item measure developed by Mendes et al. (2007) (4; 9 %), the two-item direct questions of challenge and threat (3; 7 %), the two-item appraisal measure developed by Skinner and Brewer (2002) (2; 4 %), the eight-item version of the challenge and threat construal measure (McGregor and Elliot, 2002) (2; 4 %), six affect adjectives (Folkman and Lazarus, 1985), Challenge-Threat Scale (Drach-Zahavy and Erez, 2002), four items created by Putwain et al. (2015, 2016), a single item about being challenged-threatened (Thornton et al., 2021), a single item related to resources (Lee et al., 2019), and 11-item appraisal measure developed by Mendes et al. (2007), and the Stressor Appraisal Scale (Schneider, 2008; all 1; 2 %).

Risk of bias in individual studies

We present risk of bias traffic light plots in Figure 2 and Figure 3. Most of the included studies (65.0 %) featured “low” study-level risk of bias ratings. Of the five direct-experimental studies, five (100 %) received a low risk of bias rating. Of the 57 studies that reported a correlational or indirect experimental result between CAT states and performance, 34 (61.8 %) received a low, 20 (36.4 %) an unclear, and one (1.8 %) a high risk of bias rating. We did not assess two unpublished studies. An unclear or high risk of bias rating was often associated with deviations from the typical BPSM-based measurement of CAT states and missing outcome data exceeding 20 %. The average initial interrater agreement across risk of bias domain ratings was 92.1 %.

Table 1: List of studies included in the meta-analysis

ID	Reference	N	N performance	Gender	Age	Performance type	Cardiovascular Interval			Cognitive Evaluations	Risk of Bias
							CO	TPR	CTI		
1	Arthur et al., 2019 (Study 1)	19	38	21 %	23.62	Indoor cycling	0	1	1	0	1
2	Arthur et al., 2019 (Study 2)	18	18	33 %	23.05	Indoor cycling	0	1	1	0	1
3	Baumgartner and Schneider, 2023	95	95	48 %	31.80	Math	1	1	1	0	1
4	Behnke et al., 2020	82	82	0 %	19.47	Esports	1	1	0	0	1
5	Behnke et al., 2022	241	1205	0 %	23.63	Esports	1	1	0	0	1
6	Behnke et al., 2024	300	2400	0 %	21.95	Esports	2	1	1	1	1
7	Bosshard et al., 2023	221	221	68 %	24.43	Communication	1	1	1	1	-
8	Brimmell et al., 2019	42	42	17 %	23.50	Soccer	0	1	1	1	1
9	Cabral et al., 2024	80	80	43 %	22.63	Golf	NA	0	0	0	1
10	Carenzo et al., 2020	95	159	62 %	29.90	Medical test	NA	0	0	0	1
11	Conlon et al., 2022	67	134	59.7 %	20.17	Shooting	-	0	0	0	1
12	Crowe et al., 2020	39	78	37 %	20.85	Visual search task	0	1	1	1	0
13	Dixon et al., 2020	37	37	0 %	17.95	Soccer	0	0	0	1	0
14	Gurera and Isaacowitz, 2022 (O)	47	47	47 %	72.00	Math	0	1	0	0	1
15	Gurera and Isaacowitz, 2022 (Y)	73	73	53 %	21.00	Math	0	1	0	0	1
16	Hangen et al., 2019	264	264	76 %	19.90	Math	0	1	1	1	0
17	Hase et al., 2019a	49	192	33 %	22.50	Throwing, Memory	0	1	1	1	1
18	Hase et al., 2019b	62	123	16 %	24.00	Dart	0	1	1	1	1
19	Hase et al., in preparation a	46	92	61 %	15.89	Trampoline	NA	0	0	0	1
20	Hase et al., in preparation b	33	375	9 %	23.06	Math, Memory, Throwing	0	1	1	1	1
21	Jamieson et al., 2021	478	441	60.3 %	24.16	Academic performance	-	0	0	0	1
22	Jamieson et al., 2022	339	339	62 %	24.94	Academic performance	NA	0	0	0	1
23	Jewiss et al., 2023 (Study 1)	45	38	0 %	25.00	Cricket	0	1	1	1	0
24	Jewiss et al., 2023 (Study 2)	40	320	35 %	22.00	Cricket	0	1	1	1	0
25	Jewiss et al., 2024	38	38	0 %	23.00	Cricket	0	1	1	1	0
26	Journault et al., in preparation	308	909	59 %	-	Academic performance	NA	0	0	0	1
27	Khalaf et al., 2020	107	214	61 %	24.80	Math	1	1	1	1	1
28	Laurin and Pellet, 2023	50	1400	-	20.80	Soccer	NA	0	0	0	1
29	Lee et al., 2019	499	499	52 %	14.20	Academic performance	NA	0	0	0	1
30	Malkoc et al., 2023	428	428	74 %	24.28	Academic performance	NA	0	0	0	1
31	Mansell, 2023	120	120	49 %	19.72	Golf putting	-	0	0	0	1
32	Marr et al., 2021	177	177	82 %	22.03	Memory	-	0	0	0	1
33	Miller et al., 2021a	120	120	39 %	22.22	Throwing	0	0	0	1	0
34	Miller et al., 2021b (Study 1)	80	80	30 %	21.14	Throwing	0	0	0	1	0
35	Miller et al., 2021b (Study 2)	120	120	50 %	22.62	Throwing	0	0	0	1	0
36	Moe and Putwain, 2020	60	60	50 %	18.24	Math	NA	0	0	0	1
37	Moore et al., 2018	100	100	36 %	21.94	Dart	0	1	1	1	1
38	Mosley et al., 2018	49	3675	57 %	24.10	Working memory task	NA	0	0	0	1
39	Mulvenna et al., 2023	324	324	66 %	45.40	Running	NA	0	0	0	1
40	O'Brien et al., 2020	58	58	33 %	28.00	Muscular Endurance	0	1	1	1	1
41	Petzel and Casad, 2022	140	140	100 %	21.04	Math	1	1	1	1	0
42	Sammy et al., 2017	54	54	39 %	21.72	Dart	0	1	1	1	1
43	Scheepers, 2017	104	103	100 %	21.00	Pattern recognition task	1	0	0	1	0
44	Scheepers and Keller, 2022	153	153	66 %	22.00	Tetris	1	1	1	1	1
45	Schickel et al., 2023	123	123	69 %	23.00	Presentation	NA	0	0	0	1
46	Sharpe et al., 2024a (Study 1)	28	28	11 %	22.01	Esports	NA	0	0	0	1
47	Sharpe et al., 2024a (Study 1)	33	33	12 %	21.58	Esports	NA	0	0	0	1
48	Sharpe et al., 2024a (Study 2)	57	57	33 %	19.56	Esports	NA	0	0	0	1
49	Sharpe et al., 2024b (Study 1)	44	88	0 %	21.34	Esports	-	0	0	0	1
50	Sharpe et al., 2024b (Study 2)	10	20	20 %	20.40	Esports	-	0	0	0	1
51	Simms, 2022	34	102	32 %	22.71	Vigilance task	0	1	1	1	1
52	Slater et al., 2018	83	83	30 %	20.78	Math	1	0	0	1	0
53	Smith et al., 2022	36	36	53 %	20.24	Presentation	0	1	1	1	0
54	Snijdwint and Scheepers, 2023	117	117	51 %	22.00	Cooperation comp. game	1	1	0	0	0
55	Thornton et al., 2021	120	120	48 %	22.10	Tennis	NA	0	0	0	1
56	Trotman et al., 2018	78	78	0 %	20.00	Esports	1	1	1	1	1
57	Turner et al., 2021	92	92	100 %	13.26	Netball	NA	0	0	0	1
58	van Gog et al., 2024 (Study 1)	131	131	76 %	23.81	Problem-solving tasks	-	0	0	0	1
59	van Gog et al., 2024 (Study 2)	122	122	52 %	22.89	Problem-solving tasks	-	0	0	0	1
60	van Gog et al., 2024 (Study 3)	167	167	50 %	27.28	Problem-solving tasks	-	0	0	0	1
61	van Gog et al., 2024 (Study 4)	211	211	59 %	27.29	Problem-solving tasks	-	0	0	0	1
62	Wood et al., 2018	35	35	37 %	26.00	Indoor cycling	0	1	1	1	1

Notes. N = number of participants. N performance = number of CAT-performance data pairs for the CAT variable with the most observations. Gender = percentage of Female participants. Cardiovascular Interval: 0 = before the performance, 1 = during the performance, 2 = both before and during the performance. CO = Cardiac Output. TPR = Total Peripheral Resistance. CTI = (Cardiovascular) Challenge-Threat Index. Coding for CO, TPR, CTI, and Cognitive Evaluations: 0 = not reported, 1 = reported. Risk of Bias: 1 = low, 2 = medium/unclear, 3 = high. Gurera and Isaacowitz, 2022: O = Old, Y = Young. NA = not applicable.

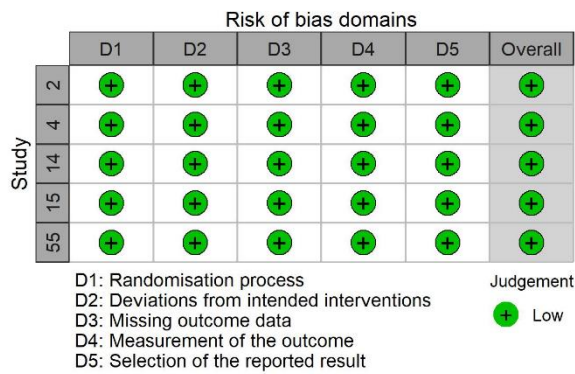


Figure 2: Traffic light plot for the ROB-2 assessment results. Note: Study ID relates to Table 1

Relationship between CAT states and performance outcomes

Overall, we found expected (positive or negative) correlations for the relationship between CAT cognitive evaluations and/or cardiovascular responses and the performance outcomes (Table 2). However, the relationship between CO reactivity and performance outcomes was not statistically significant. All mean effect sizes for the relationships should be interpreted as small. The distributions of effect sizes for the given measure are presented on Figure 4 and 5 (for more detailed figures, see Supplementary information, “CAT_Meta2024.html” file).

Furthermore, as for heterogeneity in effect sizes in the models, we did not find significant within-study (level 2) or between-study (level 3) variance. A breakdown of the total variance into the variance distributed at the three levels of the model revealed that almost all could be attributed to sampling variance. Thus, we rejected the null hypothesis of effect size homogeneity and found that the true effect size was heterogeneous and was likely to vary from effect size to effect size. This indicates that the effect sizes should not be treated as estimates of one common effect size, and thus, moderator analyses are justified to search for variables that can explain the heterogeneity of the relationship between CAT states (cognitive evaluations and/or cardiovascular responses) and performance outcomes.

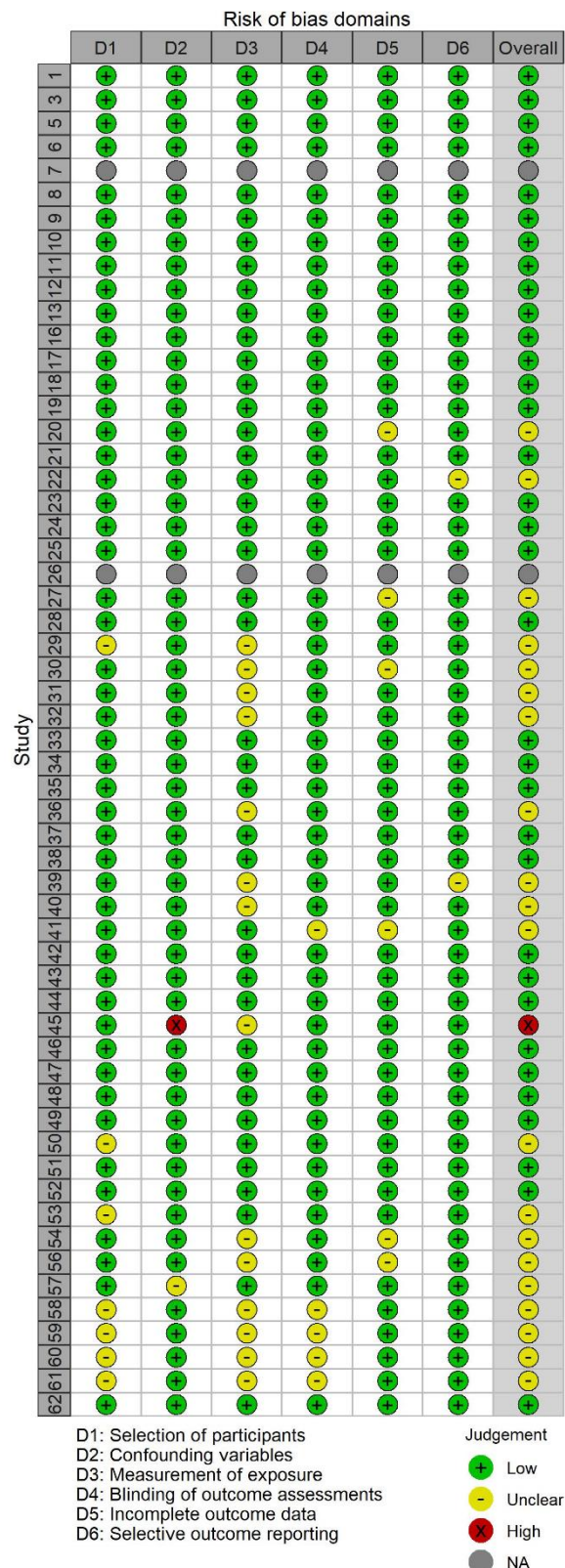


Figure 3: Traffic light plot for the ROBANS assessment results. Note: Study ID relates to Table 1

Table 2: Summary of meta-analytic review of the association between challenge and threat cardiovascular responses and/or cognitive evaluations and performance outcomes

	Pooled Effect Size					Heterogeneity			Publication bias	
	<i>k</i>	<i>r</i>	95 % <i>CI</i>	95 % <i>PI</i>	<i>Q</i>	<i>I</i> ² (Level 1)	<i>I</i> ² (Level 2)	<i>I</i> ² (Level 3)	<i>Adapted Egger test</i>	<i>Rank correlation test</i>
CO	68	0.06	-0.00, 0.13	-0.25, 0.36	151.69***	99.85	0.08	0.07	0.03	-0.05
TPR	61	-0.10**	-0.17, -0.03	-0.35, 0.17	117.86***	99.88	0.12	0.00	1.39	-0.01
CTI	73	0.10***	0.04, 0.15	-0.24, 0.42	195.52***	99.84	0.00	0.16	0.29	0.04
Evaluations	121	0.16***	0.12, 0.21	-0.13, 0.43	377.13***	99.93	0.04	0.03	1.59	-0.16*

Note. CO = Cardiac Output. TPR = Total Peripheral Resistance. CTI = (Cardiovascular) Challenge-Threat Index. *** $p < .001$, ** $p < .01$, * $p < .05$.

Publication bias

We investigated outliers by calculating studentized residuals, which identify effect sizes that disproportionately contribute to the overall heterogeneity and the results. We identified two effect sizes for CO and TPR that stood out with particularly large values (both in Petzel and Casad, 2022), with all $Z_s > 3.00$. We double-checked the effects and decided to keep them as they were clearly presented in the manuscript.

As for the publication bias assessment results, we found mixed evidence that the distribution of effect sizes was asymmetrical. First, for CO, TPR, and CTI, the trim-and-fill analysis imputed "missing effects", for the non-multilevel models (Figure 6). The analysis showed that our results might be underestimated due to the missing values and that the

'real' mean effect sizes might be bigger. The results of the rank-order correlation test and adapted Egger's regression test did not suggest an asymmetrical distribution of effects (except for CTI; see Table 2). Furthermore, we found no association between publication year and strength of CAT-performance relationship (for both CAT cardiovascular responses and cognitive evaluations), indicating that the magnitude of effect sizes did not decrease over time, meaning that more recent studies showed similar effect sizes. In sum, only two out of the five bias assessment methods applied suggested the presence of publication bias in the synthesized studies. Therefore, the evidence for publication bias affecting the results of this meta-analysis is relatively limited.

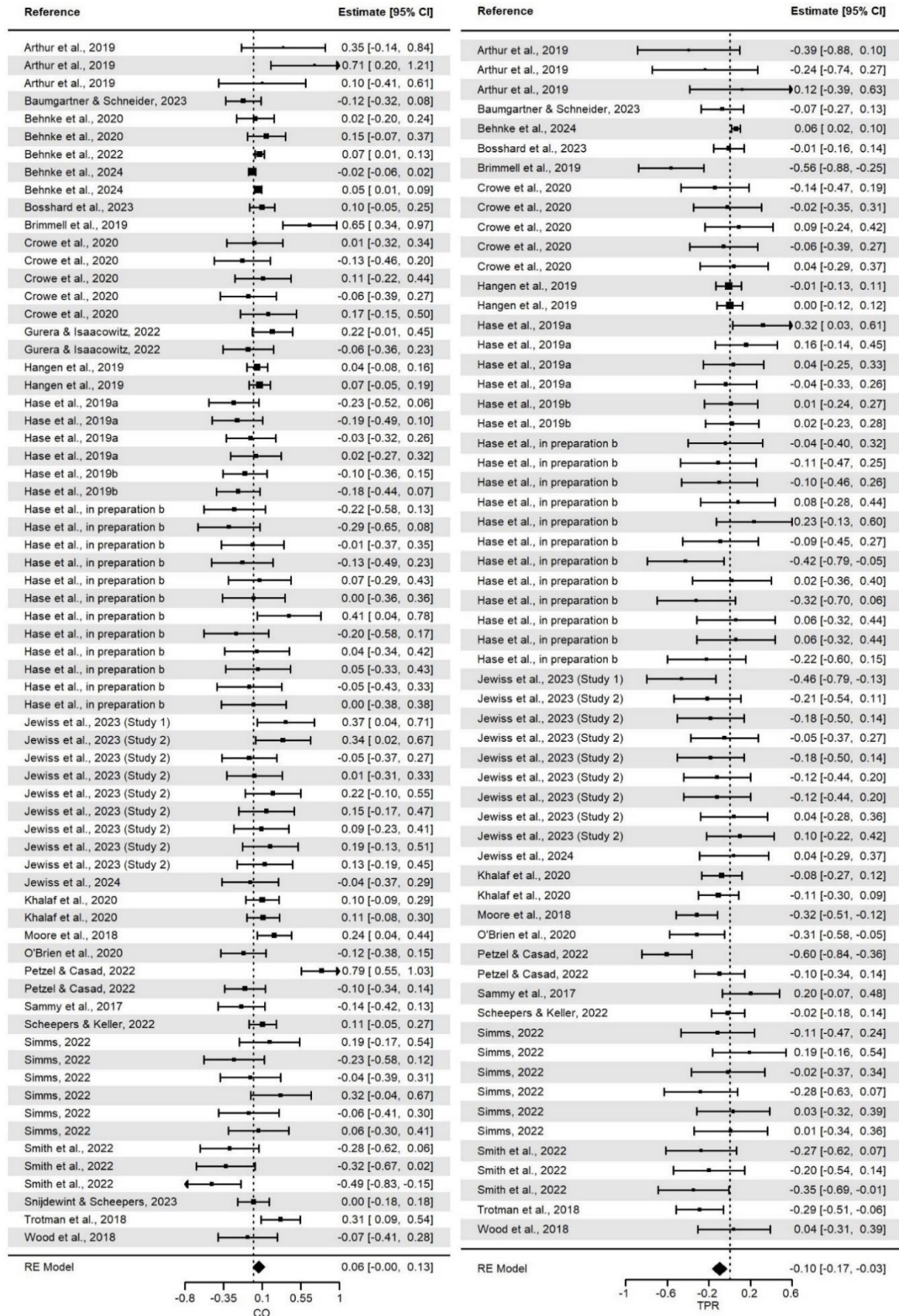


Figure 4: Forest plots for effect sizes for associations between cardiac output, total peripheral resistance and performance outcomes.

Note. CO = Cardiac Output. TPR = Total Peripheral Resistance. The square boxes represent effect sizes and sample sizes (the larger the box, the larger the sample size, contributed more to the total effect size) in each study. The lines represent 95 % confidence intervals. The diamonds represent the pooled effect size and the 95 % confidence intervals.

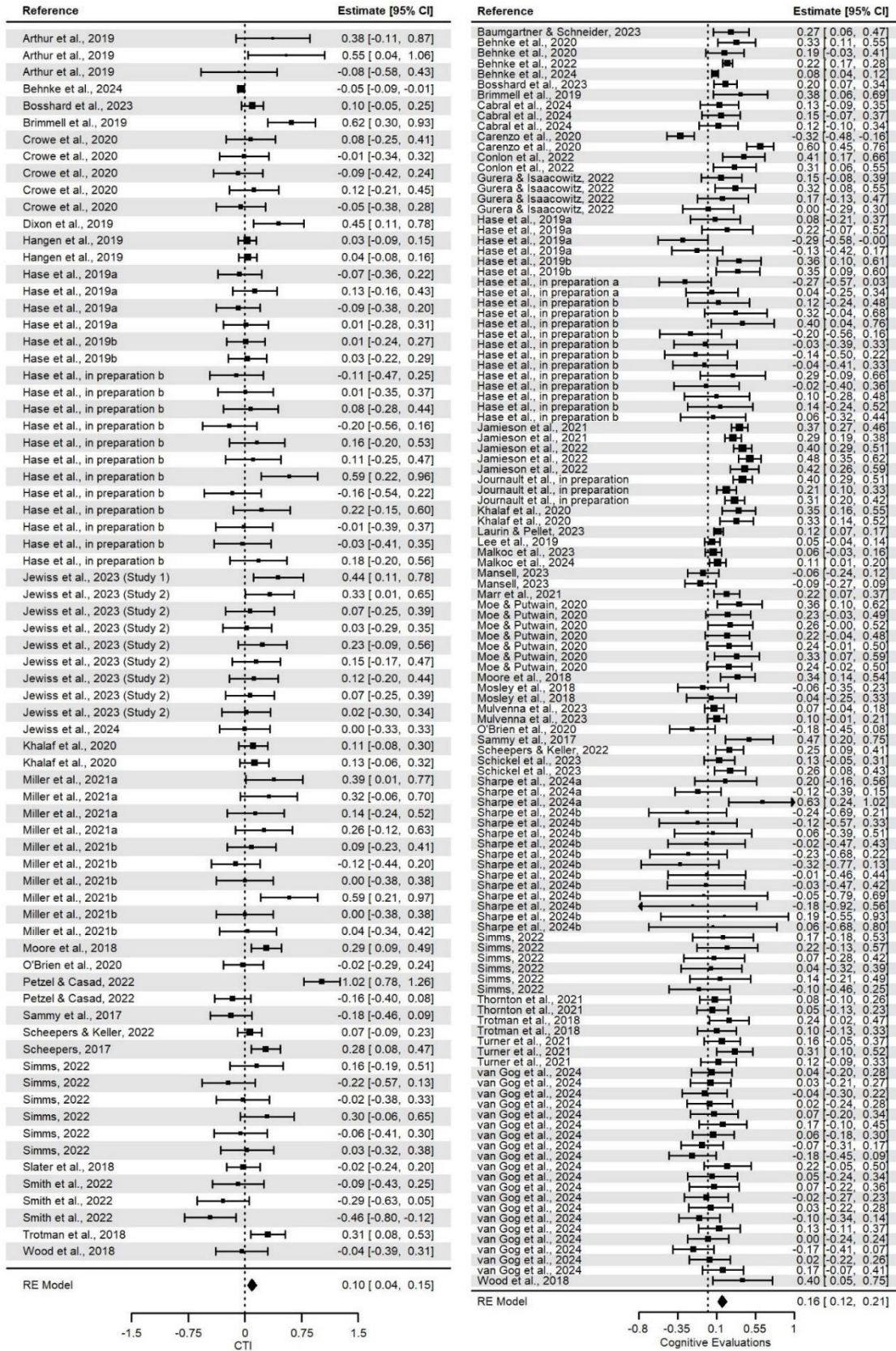


Figure 5: Forest plot for effect sizes for associations between challenge threat index, cognitive evaluations, and performance outcomes

Note. CTI = Challenge Threat Index. The square boxes represent effect sizes and sample sizes (the larger the box, the larger the sample size, contributed more to the total effect size) in each study. The lines represent 95 % confidence intervals. The diamond represents the pooled effect size and the 95 % confidence intervals.

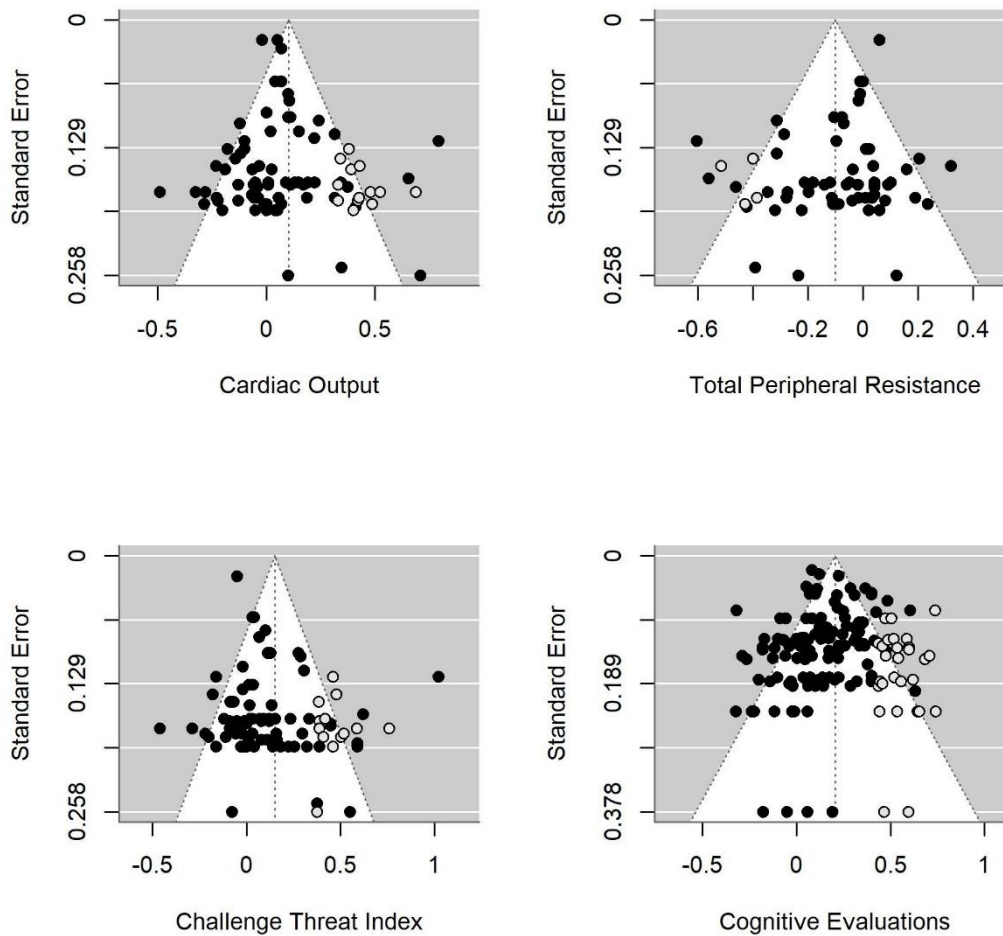


Figure 6: Funnel plots for associations between challenge and threat cardiovascular responses and cognitive evaluations and performance outcomes

Note. Black dots indicate studies included in the meta-analysis. White dots indicate studies “filled” based on the trim-and-fill method estimation.

Moderator analyses

The moderator analyses showed that under most conditions, the relationship between CAT states (cognitive evaluations and/or cardiovascular responses) and performance outcomes was not affected by most moderators. Table 3 presents the results of the omnibus tests of the moderator analyses. After adjusting for Bonferroni correction, the significant moderators were the study design for the relationship between CTI and performance outcomes, and the risk of bias category for the relationship between TPR, CTI, and cognitive evaluations and performance outcomes. When CTI was measured before the performance, the association with performance

outcomes was weaker, $r = .08$, 95 % CI [.02, .14], than when measured during performance, $r = .16$, 95 % CI [.03, .29]. Furthermore, when the risk of bias was low, the association between TPR and performance outcomes was weaker, $r = -.05$, 95 % CI [-.12, .01], than when the Risk of Bias was unclear, $r = -.25$, 95 % CI [-.38, .13]. For CAT cognitive evaluations, when the risk of bias was low, the association with performance outcomes was stronger, $r = .18$, 95 % CI [.12, .24], than when the Risk of Bias was unclear, $r = .14$, 95 % CI [.05, .21]. Finally, the difference between groups of studies for CTI ($r = .10$, vs. $r = .10$) was not interpretable.

Table 3: Results of the moderator analyses for associations between challenge and threat cardiovascular responses and cognitive evaluations and performance outcomes

	Age		Gender		Measured Before or During Performance		Publication Year		Risk of Bias	
	<i>F</i>	<i>df</i>	<i>F</i>	<i>df</i>	<i>F</i>	<i>df</i>	<i>F</i>	<i>df</i>	<i>F</i>	<i>df</i>
CO	1.18	1, 66	0.03	1, 66	2.07	2, 66	3.92	1, 84	1.54	2, 65
TPR	0.08	1, 59	0.01	1, 59	4.00*	2, 59	1.60	1, 77	9.23***	2, 58
CTI	0.02	1, 71	0.33	1, 71	6.92**	2, 71	NA	NA	5.31**	2, 70
Evaluations	0.53	1, 117	0.00	1, 118	NA	NA	1.51	1, 119	14.68***	3, 117

Note. CO = Cardiac Output. TPR = Total Peripheral Resistance. CTI = Challenge Threat Index. Results for Publication Year for CO, TPR, and CTI also include data from Behnke and Kaczmarek, 2018. NA = model did not converge. Bolded = significant after adjusting for Bonferroni correction. NA = not applicable as the models did not converge. *** $p < .001$, ** $p < .01$, * $p < .05$.

Synthesis of evidence not included in the meta-analysis

Some studies could not be included in the meta-analysis for different reasons, like a lack of reported effect sizes and/or data for the computation of effect sizes or authors' nonresponses to our data requests. Overall, we identified seven studies that could not be included in the meta-analysis for these reasons. Of these studies, some still reported results relevant to a qualitative narrative review of the CAT-performance relationship. Indeed, two studies reported null results (Chadha et al., 2023; Dixon et al., 2020; study 2), although Dixon and colleagues found a significant association between cardiovascular CAT and performance when considering player-based performance ratings. In an experimental comparison of heat and moderate temperatures in a cycling ergometer endurance test, Donnan et al. (2022) found that performance declined more in the hot condition, such that it was lower in the second half of the task. Moreover, the hot condition featured cognitive CAT evaluations more consistent with a threat state. In an experimental study, McGreary et al. (2020) found that nine out of 10 participants in a threat group did not successfully complete a cricket task, whereas 10 out of 10 participants in the challenge group successfully completed the task. Although the finding could not be included in the meta-analysis results, a chi-squared test showed

that successful performance was significantly more frequent in the challenge than threat group, $X(1) = 12.93$, $p < .001$. One study found that loss (vs. gain) framing led to increased threat cognitive evaluations (measured separately from challenge cognitive evaluations, which were unaffected by the manipulation) and, in turn, to worse working memory performance (Chen and Qu, 2023). Interestingly, no study reported a significant association contrary to the BPSM's predictions (e.g., a threat state linked to better performance outcomes than a challenge state). Finally, one article described a study protocol which had not yet led to collected data that could have been included in the meta-analysis (Guyon et al., 2020).

DISCUSSION

This systematic review and meta-analysis examined the relationship between CAT cognitive evaluations and cardiovascular responses, as conceptualized by the BPSM (Blascovich, 2008; Seery, 2011), and performance outcomes to compare the recent empirical research with that reviewed by Behnke and Kaczmarek (2018) and Hase et al. (2019c). Consistent with the previous meta-analysis (Behnke and Kaczmarek, 2018) and systematic review (Hase et al., 2019c), the present analysis established the superiority of a challenge state over a threat state on

performance outcomes when using CAT cognitive evaluations (e.g., demand resource evaluation score) and cardiovascular responses (i.e., TPR and CTI). However, the meta-analytic effects for CO did not reach statistical significance. These results, by and large, support the predictions of the BPSM and provide a relatively reliable estimate of typical effect sizes in BPSM-based CAT research (i.e., small but robust effects).

Most of the included studies featured a low risk of bias. However, some studies were flagged as unclear or high risk owing to issues such as deviations from the CAT measures typically used in research using a BPSM-based perspective (e.g., Schickel et al., 2023) or having more than 20 % missing CAT or performance outcome data in the relevant analyses (e.g., Hase et al., in prep. (b); Malkoc et al., 2023; Trotman et al., 2018). We therefore recommend that CAT researchers conduct more preregistered studies, publish the study protocols with pre-planned analyses, and publish the raw open-access data in advance to prevent publication bias and methodology-related replicability issues (see Behnke et al., 2024 for an example). However, it is worth mentioning that some of the recent studies have already implemented these recommendations; that is, they were preregistered and involved open data sharing (Bossard et al., 2023; Guyon et al., 2020; Yeager et al., 2022) or tried to replicate prior research (e.g., Jewiss et al., 2023; 2024). Furthermore, we recommend that researchers include correlation tables in their manuscripts or supplementary materials and raw data that would allow researchers to extract data for future meta-analyses. The raw data would help researchers to run analyses using innovative approaches, including accounting for non-responders (i.e., a blunted cardiovascular response; Hase et al., 2020) or using other cardiovascular measures (e.g., stroke volume instead of CO; Yeager et al., 2022).

We also examined publication bias and moderator analysis results for meta-level bias. These analyses showed mixed results. Indeed, Egger's test and Rank Correlation test did not,

but the trim-and-fill method did suggest the occurrence of publication bias. Moreover, we did not find any association between publication year and reported effect size. Thus, the effect sizes did not decrease over time, and the data did not suggest any potential replication issues. Since the trim-and-fill method did not account for the multi-level structure of the data, the risk of publication bias appears relatively small.

We also tested whether some methodological choices influenced the size of the effects. For instance, we found that measuring the CAT cardiovascular responses before or during performance was a significant moderator, but only for CTI. Unfortunately, there was a set of studies that included only CTI and did not provide data on CO and TPR separately (Dixon et al., 2020; Miller et al., 2021a, 2021b; Scheepers, 2017; Slater et al., 2018). Thus, we were unable to determine whether this methodological choice systematically influenced the strength of the association between CAT cardiovascular responses and performance outcomes or whether the additional studies including only CTI in their analyses drove this effect. Overall, the findings for CO and TPR are in line with a recent study, where researchers used multiverse analysis to examine the consistency of the effects when using different statistical and methodological approaches (Behnke et al., 2024). Furthermore, similarly to the previous meta-analysis (Behnke and Kaczmarek, 2018), age, gender, and publication year did not impact the size of the effects. Regarding publication year, the presence of the COVID-19 pandemic in the reviewed period makes the interpretation of publication year analyses more complicated when compared with previous meta-analyses (e.g., Behnke and Kaczmarek, 2018). We had hoped to test if the results of studies conducted before and after COVID-19 differed but were unable to code exact data collection dates for 91 % of studies. Thus, we recommend adding detailed study metadata (e.g., date and year of collection) in future studies to better enable comparisons like this in future meta-analyses.

Our failed attempts to explain the heterogeneity of the mean effects with our moderator analyses suggest that there might be other factors influencing the strength of the association between CAT states and performance outcomes. To differentiate their impact on effect sizes and thereby elucidate the mechanisms of the CAT-performance relationship, future research could include attentional, emotional, and behavioral processes in the study of the CAT-performance relationship (Jones et al., 2009; Vine et al., 2016). Indeed, some of the studies included in this review and meta-analysis set out to test possible mechanisms underlying the CAT-performance relationship. For instance, Brimmell et al. (2019) found that experienced soccer players who exhibited a challenge state displayed superior attentional control (e.g., longer quiet eye durations) to players who exhibited a threat state. Moreover, Wood and colleagues (2018) found that participants in a challenge state reported more positive affect, lower perceived exertion, and less self-focused attention (e.g., on movements) during a pressurized Wingate test compared to participants in a threat state. As the exact processes behind the small but robust effects of CAT on performance outcomes remain unclear, far more research on how challenge states improve performance more than threat states is required to extend the findings of this review and meta-analysis.

Our review has several practical implications, suggesting that people are less likely to succeed when they perceive their coping resources as insufficient to meet situational demands. The findings indicate that having strong confidence in one's coping resources before a motivated performance situation (e.g., exam, sporting competition) is linked to better outcomes, underscoring the importance of fostering a positive mindset about one's coping resources during training and pre-competition phases as part of comprehensive psychological preparation. Previous studies have shown that it is possible to promote challenge states through various strategies, such as encouraging task-focused instructions

(Moore et al., 2012), providing performance-related feedback (Behnke et al., 2020), using self-talk (Hase et al., 2019b), reappraising arousal (Sammy et al., 2017), applying imagery techniques (Williams et al., 2010), or engaging in pressure training (Kent et al., 2022). More recently, efforts have been made to develop interventions (e.g., arousal reappraisal; synergistic mindset) specifically designed to help individuals adopt more challenge-like mindsets (Behnke et al., 2024; Sharpe et al., 2024a; Yeager et al., 2022). We encourage researchers and practitioners to test existing interventions further and develop new ones (e.g., if-then planning; Bieleke et al., 2021), and for a systematic review and meta-analysis to summarize and synthesize the research conducted to date on interventions designed to foster a challenge (vs. threat) state, highlighting which strategies are most effective.

Strengths and limitations

The key strengths of the present systematic review and meta-analysis include the update and quantitative synthesis of relevant BPSM-based CAT research results since 2017. Specifically, the application of a meta-analysis, as well as the inclusion of both CAT cognitive evaluations and cardiovascular responses, provide advances from the previous work by Behnke and Kaczmarek (2018) and Hase et al. (2019c).

The current meta-analysis is also limited in some ways. For example, some relevant research results could not be synthesized in the meta-analysis because the publications did not report the necessary statistics, or the authors did not respond to our inquiries. Although this is not unheard of in meta-analytical research (e.g., Behnke and Kaczmarek, 2018; Behnke et al., 2022, 2023), the absence of these findings naturally implies room for improvement in the present paper's results. The absence of this data introduces a potential risk of bias in our findings, as the studies for which data could not be obtained may differ systematically from those included in the analysis (Page et al., 2022). To address this

limitation in future reviews and meta-analyses, we recommend openly sharing the code and data from the analysis. This approach would enhance transparency and allow for more comprehensive and unbiased synthesis of the evidence. Next, we observed some improvement in the range of participants' age from previous meta-analyses (e.g., Gurera and Isaacowitz, 2022; Mulvenna et al., 2023; Turner et al., 2021), still most studies have been conducted with samples of young adults, limiting the generalizability of our findings beyond this age group. To better understand the robustness of the BPSM across different stages of life, more research is needed, as cardiovascular efficiency varies with age (Mitchell et al., 2004). As a result, there is little knowledge about whether the physiological benefits associated with a challenge state also translate into improved outcomes in older adults.

Finally, the scope of the present review was restricted to a BPSM-based view of CAT states. Some research focusing on separate and independent CAT subjective measures may therefore not have been included (e.g., work underpinned by the theory of Lazarus and Folkman, 1984). Thus, it remains for future research to synthesize the findings from studies that conceptualize CAT in distinct ways and for this work to consider outcomes that are not solely performance-related (e.g., health, well-being). Indeed, research suggests that threat states, when frequently experienced, may be linked with poorer health and well-being (e.g., Brown et al., 2021; McLoughlin et al., 2024; Yeager et al., 2022).

CONCLUSION

Given the results of this systematic review and meta-analysis, we conclude that the published literature largely supports the predictions of the BPSM, with a challenge state associated with superior performance outcomes than a threat state. However, we also conclude that the field would benefit from more empirical research to test the robustness of the observed effects, especially for the separate components of CAT cardiovascular response,

where the aggregate effects did not reach statistical significance (e.g., for CO reactivity), despite being in the expected direction and the single cardiovascular CAT index reaching significance. We also conclude that the methodological quality of the studies was generally high, and the risk of publication bias was relatively small. We encourage researchers to continue investigating the predictions of the BPSM using robust methods that effectively address these publication bias and replicability issues, for example, using pre-registered studies with pre-defined analysis plans and open-access data. Given that three reviews now support the performance benefits of a challenge state relative to a threat state, further research is needed to explore the possible underlying mechanisms driving these effects (e.g., attentional, behavioral, emotional, physiological). Future work also needs to identify optimal intervention strategies for fostering a challenge state and, in turn, improving performance.

Declaration of Generative AI and AI-assisted technologies in the writing process

During the preparation of this work, the authors used Grammarly and ChatGPT4.0 to improve the readability and language of parts of the manuscript (e.g., Introduction and Discussion). After using these services, the authors reviewed and edited the content as needed and take full responsibility for the content.

Acknowledgments

The National Science Centre in Poland supported preparing this article with a research grant (UMO-2020/39/B/HS6/00685) awarded to MB. The funder had no role in research design, analysis, publishing decisions, or manuscript preparation.

Author contributions

AH: conceptualization, data curation, formal analysis, funding acquisition, investigation, methodology, project administration, supervision, visualization, writing—original draft, writing—review and editing.

MN: conceptualization, data curation, investigation, writing—original draft, writing—review and editing.

MK: data curation, investigation, writing—original draft, writing—review and editing.

KS: data curation, investigation, writing—original draft, writing—review and editing.

LM: supervision, writing—original draft, writing—review and editing.

JPJ: supervision, writing—original draft, writing—review and editing.

MB: conceptualization, data curation, formal analysis, funding acquisition, investigation, methodology, project administration, supervision, visualization, writing—original draft, writing—review and editing.

All authors gave final approval for publication and agreed to be held accountable for the work performed therein.

Data accessibility

All data and analysis code used for analysis and its output are available on the *Open Science Framework (OSF)* website: <https://osf.io/e5kz8>.

REFERENCES¹

Abdi H. The Bonferroni and Šidák corrections for multiple comparisons. In: Salkind NJ (ed). *Encyclopedia of measurement and statistics* (pp 103-7). Thousand Oaks, CA: SAGE Publications, 2007.

*Arthur TG, Wilson MR, Moore LJ, Wylie LJ, Vine SJ. Examining the effect of challenge and threat states on endurance exercise capabilities. *Psychol Sport Exerc.* 2019;44:51-9. doi: 10.1016/j.psychsport.2019.04.017.

Assink M, Wibbelink CJM. Fitting three-level meta-analytic models in R: a step-by-step tutorial. *Quant Methods Psychol.* 2016;12:154-74. doi: 10.20982/tqmp.12.3.p154.

Assink M, Spruit A, Schuts M, Lindauer R, van der Put CE, Stams GJM. The intergenerational transmission of child maltreatment: a three-level meta-analysis. *Child Abuse Negl.* 2018;84:131-45. doi: 10.1016/j.chiabu.2018.07.037.

Assink M, van der Put CE, Meeuwse MWCM, de Jong NM, Oort FJ, Stams GJJM, et al. Risk factors for child sexual abuse victimization: a meta-analytic review. *Psychol Bull.* 2019;145:459-89. doi: 10.1037/bul0000188.

*Baumgartner JN, Schneider TR. Acute biopsychosocial stress responses in mindfulness meditators and non-meditators: the mediating role of closeness. *Mindfulness.* 2023;14:1435-45. doi: 10.1007/s12671-023-02146-z.

Behnke M, Kaczmarek LD. Successful performance and cardiovascular markers of challenge and threat: a meta-analysis. *Int J Psychophysiol.* 2018;130:73-9. doi: 10.1016/j.ijpsycho.2018.04.007.

*Behnke M, Kosakowski M, Kaczmarek LD. Social challenge and threat predict performance and cardiovascular responses during competitive video gaming. *Psychol Sport Exerc.* 2020;46:101584. doi: 10.1016/j.psychsport.2019.101584.

*Behnke M, Gross JJ, Kaczmarek LD. The role of emotions in esports performance. *Emotion.* 2022;22:1059-70. doi: 10.1037/emo0000903.

Behnke M, Pietruch M, Chwiłkowska P, Wessel E, Kaczmarek LD, Assink M, et al. The undoing effect of positive emotions: a meta-analytic review. *Emotion Rev.* 2023;15(1):45-62. doi: 10.1177/17540739221104457.

*Behnke M, Lakens D, Petrova K, Chwiłkowska P, Białek SJ, Kłoskowski M, et al. Supplementary material from applying a synergistic mindsets intervention to an esports context. *The Royal Society - Collection.* 2024. doi: 10.6084/m9.figshare.c.7249302.v1.

Bieleke M, Keller L, Gollwitzer PM. If-then planning. *Eur Rev Soc Psychol.* 2021;32(1):88-122. doi: 10.1080/10463283.2020.1808936.

Blascovich J. Challenge and threat. In: Elliot AJ (ed). *Handbook of approach and avoidance motivation* (pp 431-45). New York, NY: Psychology Press, 2008.

¹ Studies included in the meta-analysis are marked with an asterisk (*).

- Blascovich J, Mendes WB. Social psychophysiology and embodiment. In: Fiske ST, Gilbert DT, Lindzey G (eds). *Handbook of social psychology*. 5th ed (pp 194-227). Hoboken, NJ: Wiley, 2010. doi: 10.1002/9780470561119.socpsy001006.
- Blascovich J, Tomaka J. The biopsychosocial model of arousal regulation. *Adv Exp Soc Psychol*. 1996;28:1-51. doi: 10.1016/S0065-2601(08)60235-X.
- Blascovich J, Seery MD, Mugridge CA, Norris RK, Weisbuch M. Predicting athletic performance from cardiovascular indexes of challenge and threat. *J Exp Soc Psychol*. 2004;40:683-88. doi: 10.1016/j.jesp.2003.10.007.
- Borenstein M, Hedges LV, Higgins JP, Rothstein HR. *Introduction to meta-analysis*. Chichester, UK: Wiley, 2009. doi: 10.1002/9780470743386.
- Borenstein M, Higgins JP, Hedges LV, Rothstein HR. Basics of meta-analysis: I2 is not an absolute measure of heterogeneity. *Res Synth Methods*. 2017;8(1):5-18. doi: 10.1002/jrsm.1230.
- *Bosshard M, Schmitz FM, Guttormsen S, Nater UM, Gomez P, Berendonk C. From threat to challenge-Improving medical students' stress response and communication skills performance through the combination of stress arousal reappraisal and preparatory worked example-based learning when breaking bad news to simulated patients: study protocol for a randomized controlled trial. *BMC Psychol*. 2023;11(1):153. doi: 10.1186/s40359-023-01167-6.
- *Brimmell J, Parker J, Wilson MR, Vine SJ, Moore LJ. Challenge and threat states, performance, and attentional control during a pressurized soccer penalty task. *Sport Exerc Perform Psychol*. 2019;8(1):63-79. doi: 10.1037/spy0000147.
- Brown DJ, Arnold R, Standage M, Turner JE, Fletcher D. The prediction of thriving in elite sport: a prospective examination of the role of psychological need satisfaction, challenge appraisal, and salivary biomarkers. *J Sci Med Sport*. 2021;24(4):373-9. doi: 10.1016/j.jsams.2020.09.019.
- Brownley KA, Hurwitz BE, Schneiderman N. Cardiovascular psychophysiology. In: Cacioppo JT, Tassinary LG, Berntson GG (eds). *Handbook of psychophysiology*. 2nd ed (pp 224-64). New York, NY: Cambridge University Press, 2000.
- *Cabral DAR, Parma JO, Morris DL, Miller MW. Conquering pressure! The effects of mild-anxiety training on motor performance under pressure during early motor learning. *J Exp Psychol Hum Percept Perform*. 2024;50:178-92. doi: 10.1037/xhp0001188.
- Cafri G, Kromrey JD, Brannick MT. A meta-meta-analysis: Empirical review of statistical power, type I error rates, effect sizes, and model selection of meta-analyses published in psychology. *Multivariate Behav Res*. 2010;45:239-70. doi: 10.1080/00273171003680187.
- *Carenzo L, Braithwaite EC, Carfagna F, Franc J, Ingrassia PL, Turner MJ, et al. Cognitive appraisals and team performance under stress: a simulation study. *Med Educ*. 2020;54:254-63. doi: 10.1111/medu.14050.
- Chadha NJ, Turner MJ, Slater MJ. Examination of cognitive appraisals, irrational beliefs, and challenge and threat evaluations in the prediction of tournament affective states and performance of competitive elite Indian golfers. *Stress Health*. 2024;40(3):e3342. doi: 10.1002/smi.3342.
- Chalabaev A, Major B, Sarrazin P, Cury F. When avoiding failure improves performance: stereotype threat and the impact of performance goals. *Motiv Emot*. 2012;36:130-42. doi: 10.1007/s11031-011-9241-x.
- Chen L, Qu L. The effects of challenge and threat states on coping flexibility: evidence from framing and exemplar priming. *Anxiety Stress Coping*. 2023;36:163-83. doi: 10.1080/10615806.2022.2059472.
- Cheung MW-L. Modeling dependent effect sizes with three-level meta-analyses: a structural equation modeling approach. *Psychol Methods*. 2014;19:211-29. doi: 10.1037/a0032968.
- *Conlon A, Arnold R, Preatoni E, Moore LJ. Pulling the trigger: The effect of a 5-minute slow diaphragmatic breathing intervention on psychophysiological stress responses and pressurized pistol shooting performance. *J Sport Exerc Psychol*. 2022;44:206-19. doi: 10.1123/jsep.2021-0213.
- *Crowe EM, Moore LJ, Harris DJ, Wilson MR, Vine SJ. In-task auditory performance-related feedback promotes cardiovascular markers of a challenge state during a pressurized task. *Anxiety Stress Coping*. 2020;33:497-510. doi: 10.1080/10615806.2020.1766681.
- Daniali H, Martinussen M, Flaten MA. A global meta-analysis of depression, anxiety, and stress before and during COVID-19. *Health Psychol*. 2023;42:124-38. doi: 10.1037/hea0001259.
- Dienstbier RA. Arousal and physiological toughness: implications for mental and physical health. *Psychol Rev*. 1989;96(1):84-100. doi: 10.1037/0033-295x.96.1.84.

- Dixon JG, Jones MV, Turner MJ. The benefits of a challenge approach on match day: Investigating cardiovascular reactivity in professional academy soccer players. *Eur J Sport Sci.* 2020;20:375-85. doi: 10.1080/17461391.2019.1629179.
- Donnan KJ, Williams EL, Stanger N. The effect of exercise-induced fatigue and heat exposure on soccer-specific decision-making during high-intensity intermittent exercise. *PLoS One.* 2022;17(12):e0279109. doi: 10.1371/journal.pone.0279109.
- Drach-Zahavy A, Erez M. Challenge versus threat effects on the goal–performance relationship. *Organ Behav Hum Decis Process.* 2002;88:667-82. doi: 10.1016/S0749-5978(02)00004-3.
- Duval S, Tweedie R. Trim and fill: A simple funnel-plot-based method of testing and adjusting for publication bias in meta-analysis. *Biometrics.* 2000;56:455-63. doi: 10.1111/j.0006-341x.2000.00455.x.
- Egger M, Davey Smith G, Schneider M, Minder C. Bias in meta-analysis detected by a simple, graphical test. *BMJ.* 1997;315(7109):629-34. doi: 10.1136/bmj.315.7109.629.
- Fernández-Castilla B, Jamshidi L, Declercq L, Beretvas SN, Onghena P, Van den Noortgate W. The application of meta-analytic (multi-level) models with multiple random effects: a systematic review. *Behav Res Methods.* 2020;52:2031-52. doi: 10.3758/s13428-020-01373-9.
- Fernández-Castilla B, Declercq L, Jamshidi L, Beretvas SN, Onghena P, Van den Noortgate W. Detecting selection bias in meta-analyses with multiple outcomes: a simulation study. *J Exp Educ.* 2021;89(1):125-44. doi: 10.1080/00220973.2019.1582470.
- Field AP, Gillett R. How to do a meta-analysis. *Br J Math Stat Psychol.* 2010;63:665-94. doi: 10.1348/000711010X502733.
- Folkman S, Lazarus RS. If it changes it must be a process: study of emotion and coping during three stages of a college examination. *J Pers Soc Psychol.* 1985;48(1):150-70. doi: 10.1037/0022-3514.48.1.150.
- *Gurera JW, Isaacowitz DM. Arousal reappraisal in younger and older adults. *Psychol Aging.* 2022;37:350-6. doi: 10.1037/pag0000674.
- Guyon AJAA, Studer RK, Hildebrandt H, Horsch A, Nater UM, Gomez P. Music performance anxiety from the challenge and threat perspective: psychophysiological and performance outcomes. *BMC Psychol.* 2020;8(1):87. doi: 10.1186/s40359-020-00448-8.
- Haddaway NR, Page MJ, Pritchard CC, McGuinness LA. PRISMA2020: an R package and Shiny app for producing PRISMA 2020-compliant flow diagrams, with interactivity for optimised digital transparency and open synthesis. *Campbell Syst Rev.* 2022;18(2):e1230. doi: 10.1002/cl2.1230.
- *Hangen EJ, Elliot AJ, Jamieson JP. Stress reappraisal during a mathematics competition: testing effects on cardiovascular approach-oriented states and exploring the moderating role of gender. *Anxiety Stress Coping.* 2019;32(1):95-108. doi: 10.1080/10615806.2018.1530049.
- *Hase A, Gorrie-Stone T, Freeman P. Tyrosine intake and cardiovascular responses in a motivated performance situation. *Sport Exerc Perform Psychol.* 2019a;8(1):80-92. doi: 10.1037/spy0000144.
- *Hase A, Hood J, Moore LJ, Freeman P. The influence of self-talk on challenge and threat states and performance. *Psychol Sport Exerc.* 2019b;45:101550. doi: 10.1016/j.psychsport.2019.101550.
- Hase A, O'Brien J, Moore LJ, Freeman P. The relationship between challenge and threat states and performance: a systematic review. *Sport Exerc Perform Psychol.* 2019c;8:123-44. doi: 10.1037/spy0000132.
- Hase A, Aan Het Rot M, de Miranda Azevedo R, Freeman P. Threat-related motivational disengagement: integrating blunted cardiovascular reactivity to stress into the biopsychosocial model of challenge and threat. *Anxiety Stress Coping.* 2020;33:355-69. doi: 10.1080/10615806.2020.1755819.
- Higgins JP, Thompson SG, Deeks JJ, Altman DG. Measuring inconsistency in meta-analyses. *BMJ.* 2003;327(7414):557-60. doi: 10.1136/bmj.327.7414.557.
- Higgins JP, Thompson SG, Spiegelhalter DJ. A re-evaluation of random-effects meta-analysis. *J R Stat Soc Series A Stat Soc.* 2009;172(1):137-59. doi: 10.1111/j.1467-985X.2008.00552.x.
- Higgins JPT, López-López JA, Becker BJ, Davies SR, Dawson S, Grimshaw JM, et al. Synthesising quantitative evidence in systematic reviews of complex health interventions. *BMJ Glob Health.* 2019;4(Suppl 1):e000858. doi: 10.1136/bmjgh-2018-000858.
- Ioannidis JP. Why most published research findings are false. *PLoS Med.* 2005;2(8):e124. doi: 10.1371/journal.pmed.0020124.
- Jamieson JP. Challenge and threat appraisals. In: Elliot AJ, Dweck CS, Yeager DS (eds). *Handbook of competence and motivation: theory and application.* 2nd ed (pp 175-91). New York, NY: Guilford Press, 2017.

- Jamieson JP, Mendes WB. Social stress facilitates risk in youths. *J Exp Psychol Gen.* 2016;145:467-85. doi: 10.1037/xge0000147.
- Jamieson JP, Valdesolo P, Peters BJ. Sympathy for the devil? The physiological and psychological effects of being an agent (and target) of dissent during intragroup conflict. *J Exp Soc Psychol.* 2014;55:221-7. doi: 10.1016/j.jesp.2014.07.011.
- Jamieson JP, Peters BJ, Greenwood EJ, Altose AJ. Reappraising stress arousal improves performance and reduces evaluation anxiety in classroom exam situations. *Soc Psychol Personal Sci.* 2016;7:579-87. doi: 10.1177/1948550616644656.
- *Jamieson JP, Black AE, Pelaia LE, Reis HT. The impact of mathematics anxiety on stress appraisals, neuroendocrine responses, and academic performance in a community college sample. *J Educ Psychol.* 2021; 113:1164-76. doi:10.1037/edu0000636.
- *Jamieson JP, Black AE, Pelaia LE, Graveling H, Gordils J, Reis HT. Reappraising stress arousal improves affective, neuroendocrine, and academic performance outcomes in community college classrooms. *J Exp Psychol Gen.* 2022;151(1):197-212. doi: 10.1037/xge0000893.
- Jefferson AL, Himali JJ, Beiser AS, Au R, Massaro JM, Seshadri S, et al. Cardiac index is associated with brain aging: the Framingham Heart Study. *Circulation.* 2010;122:690-7. doi: 10.1161/CIRCULATIONAHA.109.905091.
- *Jewiss M, Runswick OR, Greenlees I. An examination of the challenge/threat state and sport-performance relationship while controlling for past performance. *J Sport Exerc Psychol.* 2023;45:195-207. doi: 10.1123/jsep.2022-0197.
- *Jewiss M, Hodgson C, Greenlees I. Reexamining the association between preseason challenge and threat states and performance across the season. *Sport Psychol.* 2024;38:153-60. doi: 10.1123/tsp.2023-0124.
- Jones M, Meijen C, McCarthy PJ, Sheffield D. A theory of challenge and threat states in athletes. *Int Rev Sport Exerc Psychol.* 2009;2:161-80. doi: 10.1080/17509840902829331.
- Kelley DC, Siegel E, Wormwood JB. Understanding police performance under stress: insights from the biopsychosocial model of challenge and threat. *Front Psychol.* 2019;10:1800. doi: 10.3389/fpsyg.2019.01800.
- Kent S, Devonport TJ, Lane AM, Nicholls W. Implementing a pressure training program to improve decision-making and execution of skill among premier league academy soccer players. *J Appl Sport Psychol.* 2022;34:691-712. doi: 10.1080/10413200.2020.1868618.
- *Khalaf A, Nabian M, Fan M, Yin Y, Wormwood J, Siegel E, et al. Analysis of multimodal physiological signals within and between individuals to predict psychological challenge vs. threat. *Exp Syst Appl.* 2020; 140:112890. doi: 10.1016/j.eswa.2019.112890.
- Kim SY, Park JE, Lee YJ, Seo HJ, Sheen SS, Hahn S, et al. Testing a tool for assessing the risk of bias for nonrandomized studies showed moderate reliability and promising validity. *J Clin Epidemiol.* 2013;66: 408-14. doi: 10.1016/j.jclinepi.2012.09.016.
- *Laurin R, Pellet J. Affective responses mediate the body language of penalty taker–decision-making relationship from soccer goalkeepers. *Res Q Exerc Sport.* 2023;95(1):227-34. doi: 10.1080/02701367.2023.2189466.
- Lazarus R, Folkman S. *Stress, appraisal, and coping.* New York, NY: Springer, 1984.
- *Lee HY, Jamieson JP, Miu AS, Josephs RA, Yeager DS. An entity theory of intelligence predicts higher cortisol levels when high school grades are declining. *Child Dev.* 2019;90(6):e849-67. doi: 10.1111/cdev.13116.
- *Malkoc S, Macher D, Hasenhütl S, Paechter M. Good performance in difficult times? Threat and challenge as contributors to achievement emotions and academic performance during the COVID-19 outbreak. *Front Psychol.* 2023;14:1264860. doi: 10.3389/fpsyg.2023.1264860.
- *Mansell P. *Exploring how stress mindset, irrational beliefs, and stress appraisals relate to psychological wellbeing and performance under pressure [dissertation].* Birmingham, UK) University of Birmingham, 2023.
- *Marr C, Sauerland M, Otgaar H, Quaedflieg CWEM, Hope L. Mitigating the negative effects of retrieval stress on memory: an arousal reappraisal intervention. *Memory.* 2021;29:330-44. doi: 10.1080/09658211.2021.1893750.
- Matthews KA, Gump BB, Block DR, Allen MT. Does background stress heighten or dampen children's cardiovascular responses to acute stress? *Psychosom Med.* 1997;59:488-96. doi: 10.1097/00006842-199709000-00005.

- McGreary M, Eubank M, Morris R, Whitehead A. Thinking aloud: stress and coping in junior cricket batsmen during challenge and threat states. *Percept Mot Skills*. 2020;127:1095-17. doi: 10.1177/0031512520938911.
- McGregor HA, Elliot AJ. Achievement goals as predictors of achievement-relevant processes prior to task engagement. *J Educ Psychol*. 2002;94:381-95. doi: 10.1037/0022-0663.94.2.381.
- McGuinness LA, Higgins JPT. Risk-of-bias VISualization (robvis): An R package and Shiny web app for visualizing risk-of-bias assessments. *Res Synth Methods*. 2021;12(1):55-61. doi: 10.1002/jrsm.1411.
- McLoughlin E, Arnold R, Moore LJ. The tendency to appraise stressful situations as more of a threat is associated with poorer health and well-being. *Stress Health*. 2024;40(3):e3358. doi: 10.1002/smi.3358.
- Meijen C, Turner M, Jones MV, Sheffield D, McCarthy P. A theory of challenge and threat states in athletes: a revised conceptualization. *Front Psychol*. 2020;11:126. doi: 10.3389/fpsyg.2020.00126.
- Mendes WB, Park J. Neurobiological concomitants of motivational states. In: Elliot AJ (ed). *Advances in motivation science*. Vol. 1 (pp 233-70). Amsterdam: Elsevier, 2014. doi: 10.1016/bs.adms.2014.09.001.
- Mendes W, Gray HM, Mendoza-Denton R, Major B, Epel ES. Why egalitarianism might be good for your health: physiological thriving during stressful intergroup encounters. *Psychol Sci*. 2007;18:991-8. doi: 10.1111/j.1467-9280.2007.02014.x.
- *Miller AJ, Slater MJ, Turner MJ. The influence of identity leadership principles on followers' challenge and threat states and motor performance. *Psychol Sport Exerc*. 2021a;54:101909. doi: 10.1016/j.psychsport.2021.101909.
- *Miller AJ, Slater MJ, Turner MJ. Shared identity content between leader and follower influences intentional mobilization and challenge and threat states. *Psychol Sport Exerc*. 2021b;54:101914. doi: 10.1016/j.psychsport.2021.101914.
- Mitchell GF, Parise H, Benjamin EJ, Larson MG, Keyes MJ, Vita JA, et al. Changes in arterial stiffness and wave reflection with advancing age in healthy men and women: the Framingham Heart Study. *Hypertension*. 2004;43:1239-45. doi: 10.1161/01.HYP.0000128420.01881.aa.
- *Moè A, Putwain DW. An evaluative message fosters mathematics performance in male students but decreases intrinsic motivation in female students. *Educ Psychol*. 2020;40:941-60. doi: 10.1080/01443410.2020.1730767.
- Moore LJ, Vine SJ, Wilson MR, Freeman P. The effect of challenge and threat states on performance: an examination of potential mechanisms. *Psychophysiology*. 2012;49:1417-25. doi: 10.1111/j.1469-8986.2012.01449.x.
- Moore LJ, Vine SJ, Freeman P, Wilson MR. Quiet eye training promotes challenge appraisals and aids performance under elevated anxiety. *Int J Sport Exerc Psychol*. 2013a;11:169-83. doi: 10.1080/1612197X.2013.773688.
- Moore LJ, Wilson MR, Vine SJ, Coussens AH, Freeman P. Champ or chump? Challenge and threat states during pressurized competition. *J Sport Exerc Psychol*. 2013b;35:551-62. doi: 10.1123/jsep.35.6.551.
- Moore LJ, Vine SJ, Wilson MR, Freeman P. Examining the antecedents of challenge and threat states: the influence of perceived required effort and support availability. *Int J Psychophysiol*. 2014;93:267-73. doi: 10.1016/j.ijpsycho.2014.05.009.
- Moore LJ, Vine SJ, Wilson MR, Freeman P. Reappraising threat: how to optimize performance under pressure. *J Sport Exerc Psychol*. 2015;37:339-43. doi: 10.1123/jsep.2014-0186.
- *Moore LJ, Young T, Freeman P, Sarkar M. Adverse life events, cardiovascular responses, and sports performance under pressure. *Scand J Med Sci Sports*. 2018;28:340-7. doi: 10.1111/sms.12928.
- *Mosley E, Laborde S, Kavanagh E. Coping-related variables, cardiac vagal activity, and working memory performance under pressure. *Acta Psychol*. 2018;191:179-89. doi: 10.1016/j.actpsy.2018.09.007.
- *Mulvenna M, Adie JW, Tramontano C. Self-based goals, underlying reasons, performance, and discrete emotions among parkrunners. *Front Psychol*. 2023;14:1017836. doi: 10.3389/fpsyg.2023.1017836.
- *O'Brien J, Parker J, Moore L, Fryer S. Cardiovascular and cerebral hemodynamic responses to ego depletion in a pressurized sporting task. *Sport Exerc Perform Psychol*. 2020;9:183-96. doi: 10.1037/spy0000199.

- Page MJ, Higgins JPT, Sterne JAC. Assessing risk of bias due to missing evidence in a meta-analysis. In: Higgins JPT, Thomas J, Chandler J, Cumpston M, Li T, Page MJ, et al. (eds). *Cochrane handbook for systematic reviews of interventions version 6.3* (updated February 2022). Cochrane, 2022. Available from: <https://training.cochrane.org/handbook/current/chapter-13>.
- *Petzel ZW, Casad BJ. Take a chance on STEM: risk-taking buffers negative effects of stereotype threat. *J Exp Educ.* 2022;90(3):656-72. doi: 10.1080/00220973.2020.1848766.
- Putwain DW, Remedios R, Symes W. Experiencing fear appeals as a challenge or a threat influences attainment value and academic self-efficacy. *Learn Instr.* 2015;40:21-8. doi: 10.1016/j.learninstruc.2015.07.007.
- Putwain DW, Nicholson LJ, Nakhla G, Reece M, Porter B, Liversidge A. Fear appeals prior to a high-stakes examination can have a positive or negative impact on engagement depending on how the message is appraised. *Contemp Educ Psychol.* 2016;44:21-31. doi: 10.1016/j.cedpsych.2015.12.001.
- R Core Team. *R: a language and environment for statistical computing*. Vienna: R Foundation for Statistical Computing, 2023. Available from: <https://www.R-project.org/>
- Riley RD, Higgins JP, Deeks JJ. Interpretation of random effects meta-analyses. *BMJ.* 2011;342:d549. doi: 10.1136/bmj.d549.
- Sacco PL. The globalized and changing landscape of the arts: the era of post-pandemic and civil unrest. In: *The Oxford handbook of arts and cultural management* (pp C45S1-7). Oxford, UK: Oxford Univ. Press, 2023.
- *Sammy N, Anstiss PA, Moore LJ, Freeman P, Wilson MR, Vine SJ. The effects of arousal reappraisal on stress responses, performance, and attention. *Anxiety Stress Coping.* 2017;30(6):619-29. doi: 10.1080/10615806.2017.1330952.
- *Scheepers D. Intergroup status differences as challenge or threat: The role of legitimacy. *Group Process Intergroup Relat.* 2017;20(1):75-90. doi: 10.1177/1368430215595108.
- *Scheepers D, Keller J. On the physiology of flow: Bridging flow theory with the biopsychosocial model of challenge and threat. *Int J Psychophysiol.* 2022;182:119-28. doi: 10.1016/j.ijpsycho.2022.10.007.
- *Schickel M, Minkley N, Ringeisen T. Performance during presentations: a question of challenge and threat responses? *Contemp Educ Psychol.* 2023;73:102168. doi: 10.1016/j.cedpsych.2023.102168.
- Schmidt FL, Hunter JE. *Methods of meta-analysis. Correcting error and bias in research findings*. Thousand Oaks, CA: SAGE Publ., 2015. doi: 10.4135/9781483398105.
- Schneider TR. The role of neuroticism on psychological and physiological stress responses. *J Exp Soc Psychol.* 2004;40:795-804. doi: 10.1016/j.jesp.2004.04.005.
- Schneider TR. Evaluations of stressful transactions: what's in an appraisal? *Stress Health.* 2008;24:151-8. doi: 10.1002/smi.1176.
- Schooler J. Unpublished results hide the decline effect. *Nature.* 2011;470(7335):437. doi: 10.1038/470437a.
- Seery MD. Challenge or threat? Cardiovascular indexes of resilience and vulnerability to potential stress in humans. *Neurosci Biobehav Rev.* 2011;35:1603-10. doi: 10.1016/j.neubiorev.2011.03.003.
- Seery MD. The biopsychosocial model of challenge and threat: using the heart to measure the mind. *Soc Personal Psychol Compass.* 2013;7:637-53. doi: 10.1111/spc3.12052.
- Seery MD, Weisbuch M, Hetenyi MA, Blascovich J. Cardiovascular measures independently predict performance in a university course. *Psychophysiology.* 2010;47:535-9. doi: 10.1111/j.1469-8986.2009.00945.x.
- *Sharpe BT, Obine EAC, Birch PDJ, Pocock C, Moore LJ. Performance breakdown under pressure among esports competitors. *Sport Exerc Perform Psychol.* 2024a;13(1):89-109. doi: 10.1037/spy0000337.
- *Sharpe BT, Leis O, Moore L, Sharpe AT, Seymour S, Obine EA, et al. Reappraisal and mindset interventions on pressurised esports performance. *Appl Psychol.* 2024b;73:2178-99. doi: 10.1111/apps.12544.
- *Simms M. *Organizational stressors, acute illness symptoms, and performance among athletes* [dissertation]. Bath, UK: University of Bath, 2022.
- Skinner N, Brewer N. The dynamics of threat and challenge appraisals prior to stressful achievement events. *J Pers Soc Psychol.* 2002;83:678-92. doi: 10.1037//0022-3514.83.3.678.
- *Slater MJ, Turner MJ, Evans AL, Jones MV. Capturing hearts and minds: the influence of relational identification with the leader on followers' mobilization and cardiovascular reactivity. *Leadersh Q.* 2018;29:379-88. doi: 10.1016/j.leaqua.2017.08.003.

- *Smith M, Runswick OR, Jewiss M, Birch P. Examining the theory of challenge and threat states in athletes: do predictions extend to academic performance? *Educ Psychol.* 2022;42:1127-44. doi: 10.1080/01443410.2022.2069229.
- *Snijdwint JA, Scheepers D. Group-based flow: the influence of cardiovascular synchronization and identifiability. *Psychophysiology.* 2023;60(5):e14227. doi: 10.1111/psyp.14227.
- Sterne JAC, Savović J, Page MJ, Elbers RG, Blencowe NS, Boutron I, et al. RoB 2: a revised tool for assessing risk of bias in randomised trials. *BMJ.* 2019;366:4898. doi: 10.1136/bmj.l4898.
- *Thornton C, Sheffield D, Baird A. Exposure to contact sports results in maintained performance during experimental pain. *J Pain.* 2021;22(1):68-75. doi: 10.1016/j.jpain.2020.03.008.
- Tomaka J, Blascovich J, Kelsey RM, Leitten CL. Subjective, physiological, and behavioral effects of threat and challenge appraisal. *J Pers Soc Psychol.* 1993;65:248-60. doi: 10.1037/0022-3514.65.2.248.
- *Trotman GP, Williams SE, Quinton ML, van Zanten JJV. Challenge and threat states: examining cardiovascular, cognitive, and affective responses to two distinct laboratory stress tasks. *Int J Psychophysiol.* 2018;126:42-51. doi: 10.1016/j.ijpsycho.2018.02.004.
- *Turner MJ, Massie R, Slater MJ, Braithwaite E. Do challenge and threat evaluations predict netball performance and selection at trials in youth netball players? *Sport Exerc Perform Psychol.* 2021;10(1):71-87. doi: 10.1037/spy0000248.
- Uphill MA, Rossato CJL, Swain J, O'Driscoll J. Challenge and threat: a critical review of the literature and an alternative conceptualization. *Front Psychol.* 2019;10:1255. doi: 10.3389/fpsyg.2019.01255.
- Van den Noortgate W, López-López JA, Marín-Martínez F, Sánchez-Meca J. Three-level meta-analysis of dependent effect sizes. *Behav Res Methods.* 2013;45:576-94. doi: 10.3758/s13428-012-0261-6.
- *van Gog T, Janssen E, Lucas F, Taheij M. A motivational perspective on (anticipated) mental effort investment: the biopsychosocial model of challenge and threat. *Educ Psychol Rev.* 2024;36(2):54. doi: 10.1007/s10648-024-09861-6.
- Viechtbauer W. Conducting meta-analyses in R with the metafor package. *J Stat Softw.* 2010;36(3):1-48. doi: 10.18637/jss.v036.i03.
- Vine SJ, Freeman P, Moore LJ, Chandra-Ramanan R, Wilson MR. Evaluating stress as a challenge is associated with superior attentional control and motor skill performance: testing the predictions of the biopsychosocial model of challenge and threat. *J Exp Psychol Appl.* 2013;19:185-94. doi: 10.1037/a0034106.
- Vine SJ, Uiga L, Lavric A, Moore LJ, Tsaneva-Atanasova K, Wilson MR. Individual reactions to stress predict performance during a critical aviation incident. *Anxiety Stress Coping.* 2015;28:467-77. doi: 10.1080/10615806.2014.986722.
- Vine SJ, Moore LJ, Wilson MR. An integrative framework of stress, attention, and visuomotor performance. *Front Psychol.* 2016;7:1671. doi: 10.3389/fpsyg.2016.01671.
- Williams SE, Cumming J, Balanos GM. The use of imagery to manipulate challenge and threat appraisal States in athletes. *J Sport Exerc Psychol.* 2010;32:339-58. doi: 10.1123/jsep.32.3.339.
- *Wood N, Parker J, Freeman P, Black M, Moore L. The relationship between challenge and threat states and anaerobic power, core affect, perceived exertion, and self-focused attention during a competitive sprint cycling task. *Prog Brain Res.* 2018;240:1-17. doi: 10.1016/bs.pbr.2018
- Yeager DS, Bryan CJ, Gross JJ, Murray J, Krettek Cobb D, et al. A synergistic mindsets intervention protects adolescents from stress. *Nature.* 2022;607(7919):512-20. doi: 10.1038/s41586-022-04907-7.