Challenge and threat states: An examination of variance components, interventions, and performance

A. Hase

A thesis submitted for the degree of Doctor of Philosophy in Sport and Exercise Psychology

School of Sport, Rehabilitation and Exercise Sciences

University of Essex

Date of final submission July 2019

Abstract

This thesis reports on a systematic review and four empirical studies that examined cognitive and cardiovascular indicators of challenge and threat states as specified by the biopsychosocial model (Blascovich, 2008a). The overarching research aim was to provide novel insights into challenge and threat states including their variability between persons and situations, their responsiveness to interventions, and their relationship with performance. The systematic review found that a challenge state was associated with better performance than a threat state across various outcomes and research designs, indicating that the analysis of challenge and threat states may provide useful information for sport psychologists. The first two empirical studies indicated that challenge and threat states vary largely as a function of personal and person by situation interactional components, suggesting that these factors are promising targets for interventions. The next two studies showed that three established interventions for optimising performance by targeting psychological function (instructional and motivational self-talk) and physiological function (tyrosine supplement) did not promote a challenge state, but instead modified the relationship between challenge and threat states and performance. Precisely, tyrosine and instructional self-talk decreased performance differences between challenge and threat states, indicating a potential for helping athletes in a threat state. The interrelationships between cognitive and cardiovascular indicators of challenge and threat states and performance in the present research were inconsistent with those predicted by the biopsychosocial model, provoking discussion about the applicability of its predictions. Future research directions include conducting systematic reviews on outcomes other than performance; conducting research into dispositional variables and person by situation interactions to reveal new correlates of challenge and threat states; examining how interventions affect challenge and threat states and their relationship with performance; and more closely examining (moderators of) the interrelationships between cognitive challenge and threat evaluations, cardiovascular challenge and threat responses, and performance.

Table of Contents

Acknowledgements

This thesis is the product of countless hours of planning and discussion, literature review, data collection, data analysis, writing, proofreading, and revising. I would like to thank Paul Freeman, my supervisor, for guiding me through every step of the process, making sure I always knew what to expect and what was expected from me, always being open to my ideas and perspectives, and being enthusiastic about sharing his visions with me. Next to his professionalism, I would also like to thank Paul for being an extremely decent person who never made me feel unwelcome in his office. However, even with his help, there would be no thesis without the countless participants who agreed to take part in my studies. There were over a hundred of you, and I am grateful for every single one who took part. I would also like to thank my collaborators (Lee Moore, Inge Korte, Melvin Dokter, Lieke Schiphof-Godart, Tyler Gorrie-Stone, Di Fu, Stevie Farrow, Scott Wright, Jacob Hood); and supervisory board members for their constructive input in the numerous stages of my PhD period at the University of Essex. I would also like to extend my thanks to the School of Sport, Rehabilitation and Exercise Sciences and the School of Biological Sciences as they have provided a great academic environment in which I learned a lot of new things over the last four years and met many great academics. The University of Essex more generally provided a very welcoming environment to me as an international student and allowed me to meet many like-minded people from all over the world. Lastly, I would like to thank my family and friends for supporting me before and throughout my time at Essex. I appreciate that I might not even have had the opportunity to apply for this PhD project without the privileges that they endowed me with, and I am deeply grateful for that.

Next to all of my family members and friends who are able to witness this thesis, I dedicate this thesis to my grandfather Hermann Hackstein (1935-2012), who had a tremendous influence on both my academic and my personal development, which I will always try to remember.

List of Tables

List of Figures

Preface

Part of this PhD thesis has been published in peer-reviewed academic journals. I have included the published chapters as they appear in their respective journal articles. Therefore, there may be overlap in content between the published and the remaining chapters. To clarify which chapters have been published, I have added a note and the respective citation on the first page of each published chapter.

Chapter 1

General Introduction to Relevant Challenge and Threat Theory and Research

Life is filled with self-relevant and potentially stressful situations that require people to perform under evaluative pressure. Most people have experienced such situations in school exams, job interviews, and/or sport competitions. However, these situations may not always feel the same. Sometimes, a person may feel as if they are thriving under the pressure, whereas other times, they may feel as if they are threatened by the situation and negative consequences are to strike them. The biopsychosocial model (BPSM, Blascovich, 2008a; Blascovich & Mendes, 2000) of challenge and threat (CAT) is a widely used psychophysiological framework that picks up this idea of important situations being evaluated as challenging or threatening. It is rooted in two popular strands of research, namely Lazarus and Folkman's cognitive appraisal theory (Lazarus & Folkman, 1984; Lazarus, 1999) and Dienstbier's research on physiological toughness (Dienstbier, 1989).

Cognitive appraisal theory (Lazarus & Folkman, 1984), which inspired the social psychological component of the BPSM, holds the core assumption that cognitive appraisal processes are an immediate reaction to the perception and subsequent experience of a potentially stressful situation. Cognitive appraisal theory specifies that primary and secondary appraisals determine the stressfulness and the stress appraisal of the given situation, which in turn is predicted to influence emotional, physiological, and behavioural responses to the situation. In this theoretical framework, primary appraisal represents what an individual believes to be at stake; for example, a personal goal or core values (Lazarus, 1999). If there is nothing at stake, then no stress should ensue. If something important is at stake, cognitive appraisal theory posits that four different alternatives may result: a challenge, threat, harm/loss, and benefit appraisal. The latter two only represent post-hoc appraisals of a stressful situation in which a negative (i.e., harm or loss) or a positive (i.e., benefit) outcome was already experienced, and are not of interest here. On the contrary, both CAT appraisals represent anticipatory appraisals of a stressful situation in cognitive appraisal theory. Secondary appraisal represents the evaluation of personal coping options to deal with the situation (Lazarus, 1999), which is particularly relevant in case of a challenge, threat, or harm/loss appraisal. Cognitive appraisal theory also holds that cognitive appraisals are dynamic and conscious processes and based on on-line information from the stressful situation, and thus may change as the situation proceeds. Cognitive appraisal theory has been widely adopted by sport psychological researchers, who have used CAT appraisals to predict outcomes like performance, mental toughness, and emotions in sport contexts (e.g., Freeman & Rees, 2009; Levy, Nicholls, & Polman, 2012; Skinner & Brewer, 2004). One of the key features of cognitive appraisal theory that the BPSM disagrees with is that CAT appraisals can occur at the same time (e.g., when there is something to gain, but also something to lose).

The physiological influence in the BPSM is based on Dienstbier's (1989) model of physiological toughness, which sought to change the idea that sympathetic nervous system arousal is generally counterproductive in dealing with intermittent stressors. In particular, Dienstbier proposed an effort mobilisation pattern of *physiological toughness* characterised by sympathetic-adrenomedullary arousal, which he proposed produces resistance to brain catecholamine depletion and suppressed pituitary-adrenocortical responses when exposed to stressors. This in turn should lead to energy mobilisation (mediated by increased blood flow, blood glucose levels, and peripheral vasodilation), good performance even in complex tasks, and immune system maintenance (i.e., a positive health outcome). Dienstbier proposed that the idea of physiological toughness could present a viable alternative to the view that adaptive coping with stress should always involve arousal reduction. However, Dienstbier did not completely abandon the idea of maladaptive stress, as he described the effort mobilisation pattern of *physiological weakness* as the maladaptive counterpart to physiological toughness. This pattern involved pituitary-adrenocortical arousal, which stimulates adrenocorticotropic hormone and ultimately cortisol release. According to the model, this leads to tension, poor performance, and immune suppression (i.e., a harmful health outcome). The patterns of physiological toughness and weakness served as templates for the challenge and the threat state, respectively.

Starting in the early 1990s, Blascovich, Tomaka and colleagues started to research how Lazarus and Folkman's cognitive appraisals of CAT related to Dienstbier's effort mobilisation patterns of physiological toughness and weakness (Tomaka, Blascovich, Kelsey, & Leitten, 1993). Over the next two decades, the resultant extension and amalgamation of these previous theoretical models produced the BPSM (Blascovich, 2008a). This chapter introduces the BPSM as well as related topics that are not (or only partly) covered in subsequent chapters. For example, it will discuss related theoretical models, antecedents to and outcomes of CAT states, challenges to the BPSM, and existing gaps in the literature.

1.1 The BPSM of CAT

The BPSM of CAT (Blascovich, 2008a; Blascovich & Mendes, 2000) describes a unidimensional continuum of stress responses ranging from a relatively adaptive (the challenge state) to a relatively maladaptive response (the threat state). This continuum occurs in the context of motivated performance situations and is conceptualised in terms of differences in cognitive demand and resource evaluations and cardiovascular responses. Motivated performance situations (e.g., sport competitions, academic exams, or job interviews) are goal-relevant, evaluative, and potentially stressful, requiring adequate active performance in order to ensure personal wellbeing and growth (Blascovich & Mendes, 2000). The BPSM predicts that in the context of a motivated performance situation, and assuming task engagement, cognitive evaluations of personal coping resources and situational demands trigger cardiovascular responses distinguishing CAT states on the physiological level. It also predicts that these evaluations are subject to continuous reappraisal as new information is obtained and therefore CAT states may vary during the situation (Blascovich, 2008a).

The BPSM (Blascovich, 2008a) operationalises the prerequisite of psychological task engagement on the cardiovascular level as increases in heart rate (HR) and ventricular contractility (VC). Assuming task engagement, a challenge state is characterised by the largely subconscious evaluation that one's personal coping resources match or exceed situational demands. Physiologically, a challenge state is marked by relative vasodilation and an increase in cardiac performance, marked by increases HR, VC, and cardiac output (CO), and a decrease in total peripheral resistance (TPR; Tomaka et al., 1993). These effects are thought to be due to sympathetic-adrenomedullary activation, which involves the release of the catecholamine neurotransmitters epinephrine and norepinephrine. In contrast, a threat state is characterised by an evaluation that coping resources fall short of situational demands. This entails relative vasoconstriction and a lesser cardiac performance increase than in a challenge state, marked by increases in HR and VC, little change in CO, and little change or increases in TPR. The BPSM attributes the threat-related cardiovascular response pattern to hypothalamic-pituitary-adrenal axis activity that counteracts the sympathetic-adrenomedullary effects via cortisol release (Blascovich, 2008a). Correlational validation research supported the prediction that a challenge evaluation (resources outweighing demands) is associated with relatively greater CO and HR, and relatively lower TPR than a threat evaluation (demands outweighing resources; Tomaka et al., 1993). Furthermore, experimental validation studies supported the prediction that cognitive CAT evaluations triggered cardiovascular responses, not vice versa (Tomaka, Blascovich, Kibler, & Ernst, 1997). On the one hand, cardiovascular CAT responses in line with the BPSM's predictions were elicited with CAT scripts designed to elicit a challenge or a threat state in anticipation of a laboratory-based mental arithmetic task. On the other hand, when challenge- or threat-like cardiovascular responses were elicited via a physiological approach (physical exercise versus no exercise to elicit relative challenge in one study and cold versus warm water immersion to elicit relative threat in another study), no differences in cognitive evaluations of personal coping resources and situational demands were observed. Figure 1.1 illustrates the main predictions of the BPSM.

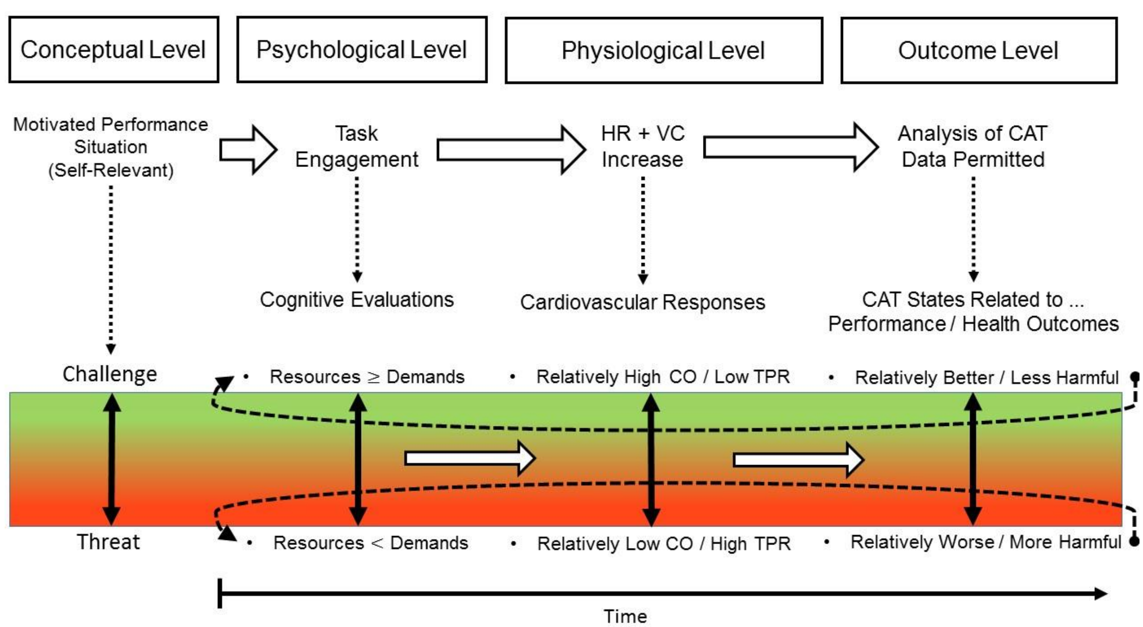

Figure 1.1. Main predictions of the BPSM.

1.2 Antecedents to CAT States

Blascovich (2008a) specified the psychological antecedents to cognitive CAT evaluations in the BPSM (i.e., perceived coping resources and situational demands) as bipolar continua that go beyond simple task characteristics, and may be interactive and interrelated. The list of antecedents includes the continua of safety-danger, uncertainty-certainty, novelty-familiarity, low-high required effort, poor-excellent skills or abilities, low-high knowledge, no-many present others (i.e., spectators or evaluators), weak-strong attitudes, and weak-strong beliefs. An example of the interrelated and interactive nature of CAT antecedents could be when the task is novel, as novelty may also provoke uncertainty regarding the outcome of the task. This could in turn increase evaluated situational demands. Another example would be when an individual possesses excellent task-relevant skills, and therefore also has relatively high certainty of succeeding, thereby improving resource evaluations. The BPSM also specifies that CAT antecedents may influence both demand and resource evaluations at the same time.

The BPSM's predictions regarding antecedents have been tested in some previous studies. For example, Moore and colleagues manipulated perceived required effort and available support and found that participants who were led to believe that the upcoming task required low effort rated cognitive evaluations more consistent with a challenge state, exhibited a cardiovascular response more consistent with a challenge state, and outperformed the low required effort group (Moore, Vine, Wilson, & Freeman, 2014). In contrast, support availability did not influence CAT states. Also, Seery, Weisbuch, and Blascovich (2009) framed task instructions in a way that instructed participants to try to gain money, or avoid losing money (i.e., consistent with an approach or avoidance goal orientation) and observed cardiovascular responses more consistent with a challenge state in the gain than in the loss framing group. The prediction that CAT antecedents may influence both demand and resource evaluations was also tested by Moore and colleagues (2014). They found that manipulating perceived required effort led to cognitive evaluations of more resources and less demands in the low, relative to the high required effort condition, but no such finding was observed for high versus low support availability. However, many of the antecedent continua specified by the BPSM remain to be tested regarding this prediction.

It should also be noted that the BPSM acknowledges that it does not provide an exhaustive list of CAT antecedents (Blascovich, 2008a). For instance, physical health (e.g., cold symptoms) could influence both resource and demand evaluations. Even the appraisal of physical symptoms may influence CAT states (Moore, Vine, Wilson, & Freeman, 2015). When Moore and colleagues gave arousal reappraisal or control instructions to two groups, they found that reappraising arousal as helpful for performance led to a cardiovascular response less reflective of a threat state and better performance than the control instructions.

1.3 Outcomes of CAT States

1.3.1 Performance

The CAT continuum is predicted to influence performance, with a challenge state being related to better performance than a threat state (Blascovich, 2008a; Jones et al., 2009). This prediction has been supported in various studies (e.g., Blascovich, Seery, Mugridge, Norris, & Weisbuch, 2004; Scholl, Moeller, Scheepers, Nuerk, & Sassenberg, 2017; Vine et al., 2015), but prior to this thesis, the field lacked a systematic review of all studies examining the relationship between CAT states conceptualised in consistency with the BPSM and performance. Thus, the BPSM's prediction is examined in chapter 2, which reports on a systematic literature review on this topic. Coincidentally, an independent group also recognised the need for a literature review and conducted a meta-analysis of 19 studies examining the relationship between cardiovascular CAT measures and performance (Behnke & Kaczmarek, 2018). Although the meta-analysis found evidence of weaker effects being underrepresented in the literature, the association between cardiovascular CAT states and performance was supported after controlling for this bias. Thus, Behnke and Kaczmarek concluded that the literature supported the validity of the BPSM in the prediction of performance.

Though better performance in a challenge than in a threat state appears to be the rule, there are some exceptions. Blascovich (2008a) highlighted that individuals in a threat state appear to perform better on vigilance tasks (Hunter, 2001), and some research found threat to be associated with better performance than challenge on an information-integration task (i.e., a form of learning based on procedural-based working memory system; Ell, Cosley, & McCoy, 2011). This might to be due to competition for cognitive resources between hypothesis-testing and information-integration systems. Thus, a threat state may have negative effects on the hypothesis-testing system, reducing this aspect of cognitive performance, but this may free up more cognitive resources for information-integration processes (Ell et al., 2011). Finally, in a series of experiments, Feinberg and Aiello (2010, study 1) found that the threat group outperformed the challenge group on a mental arithmetic task, although the manipulation check was only marginally significant and not based on cardiovascular indices of CAT.

1.3.2 Mental and Physical Health

There are several other key outcomes associated with CAT states. For example, Blascovich (2008b) predicted that in the long run, a threat state might have deleterious health effects, whereas a challenge state is a benign, or even health-improving physiological pattern. In particular, Blascovich predicted that repetitive threat state experience can lead to ischemic heart disease and hypertension, because the cardiovascular pattern of a threat state (relative vasoconstriction and higher VC) puts strain on the vasculature, especially the coronary arteries. Furthermore, he predicted that repetitive threat state experience and the related release of cortisol would worsen immune system function (e.g., due to the immunosuppressant long-term effect of cortisol). Blascovich also predicted that a threat state may indirectly deteriorate health, for example, by increasing an individual's risk of experiencing anxiety and depression. The cognitive biases toward threatening stimuli in anxiety, depression, and post-traumatic stress disorder support this view (e.g., Beck & Clark, 1988; Armony, Corbo, Clément, & Brunet, 2005), although it might also be that mental disorders or trauma increase the likelihood of experiencing a threat state.

1.3.3 Other Outcomes

Moore and colleagues conducted two studies in which they examined attentional and behavioural aspects of golf putting performance next to objective performance measures (i.e., putts holed, mean radial error) and found that challenged participants displayed more efficient gaze behaviour, less conscious processing, longer quiet eye durations, more efficient putting kinematics, and muscle activity than threatened participants (Moore, Vine, Wilson, & Freeman, 2012; study 2, Moore, Wilson, Vine, Coussens, & Freeman, 2013). A similar study examined CAT evaluations and cardiovascular responses in a football penalty shooting task. Its results indicated that a challenge, relative to a threat cardiovascular response, was associated with longer quiet eye durations and lower search rates (Brimmel, Parker, Wilson, Vine, & Moore, 2019).

A study by Vine and colleagues found that CAT evaluations predicted target locking in their outcome task, which represented an indicator of attentional function (Vine, Freeman, Moore, Chandra-Ramanan, & Wilson, 2013). In particular, cognitive evaluations consistent with a challenge state were associated with greater target locking than those consistent with a threat state. Furthermore, one study examined how CAT states related to nonverbal behaviour in a football penalty shooting task (Brimmel, Parker, Furley, & Moore, 2018). Participants rated demands and resources and provided cardiovascular data prior to performing, and their subsequent performance was filmed and judged by 71 untrained observers. The results showed that evaluations consistent with a challenge state were related to higher ratings of dominance, confidence, composure, challenge, and competence than evaluations consistent with a threat state, but no associations were found on CAT variables. Taken together, these results indicate more efficient attentional and behavioural functioning in a challenge, relative to a threat state.

1.4 Challenges to the BPSM

Next to the empirical tests of the BPSM's predictions presented above and in chapter 2, the BPSM has also been challenged on a theoretical level. In 2003, Wright and Kirby published a critical commentary of the BPSM and associated literature. Wright and Kirby criticised the BPSM perspective on situational demands and how CAT states result from demand and resource evaluations, goal-relevance and task engagement in motivated performance situations, and cardiovascular predictions. In the same issue, Blascovich and colleagues responded to Wright and Kirby's criticism. They argued that Wright and Kirby's account was based on a misunderstanding and selective discussion of the BPSM, and on a rational-economic approach that is inconsistent with the social psychological approach guiding the BPSM (Blascovich, Mendes, Tomaka, Salomon, & Seery, 2003). The key points of this exchange are detailed below.

Wright and Kirby's (2003) first criticism was that the BPSM does not precisely specify (i.e., in an objectively testable way) how required effort, uncertainty, and danger interact to produce demand evaluations. In response, Blascovich and colleagues (2003) argued that Wright and Kirby misrepresented the BPSM, portraying it as a rational-economic theory rather than a social psychological theory. For example, the BPSM clearly states that cardiovascular measurement of CAT variables is preferable over self-reports of cognitive evaluations, because the latter may not always be accurate. This lack of accuracy may be due to demand and resource evaluations not always being conscious or being distorted by non-conscious biases (e.g., self-presentation concerns; Blascovich & Mendes, 2000).

Wright and Kirby's (2003) second critique claimed that the conceptualisation of situational demands in the BPSM was flawed and that CAT states did not result from objectively testable balances of demands and resources. Blascovich and colleagues argued that Wright and Kirby's critique was not valid because not all antecedents to demand and resource evaluations reflect rational economic calculations, and neither do their outcomes (i.e., whether an individual experiences a challenge versus a threat state). Indeed, the research of Kruger and Dunning (1999) is a very good example of a serious bias in the conscious evaluation of skills and abilities, which in turn influence CAT states in the BPSM. An example of this would be the case of someone with very low cognitive ability evaluating a mental arithmetic task as a challenge – not because of the actual difficulty of the task, but because of their inability to make a rational calculation of their own abilities, the task's difficulty level, or both. Furthermore, Blascovich and colleagues supported their rebuttal with some of their own work showing subconscious influences in CAT states, such as the ethnicity or socioeconomic status of a partner in a cooperative word-finding task influencing cardiovascular CAT responses (e.g., Mendes, Blascovich, Lickel, & Hunter, 2002). It should be added that since the 2003 debate, Blascovich has re-conceptualised the nature of antecedents to demand and resource evaluations as reflecting single continua (i.e., one antecedent variable, such as danger-safety, can influence both demand and resource evaluations; Blascovich, 2008a).

Wright and Kirby (2003) also claimed that the terminology with which Blascovich and colleagues defined self-relevance in motivated performance situations was vague across publications (e.g., Blascovich & Tomaka, 1996; Blascovich & Mendes, 2000). Precisely, they criticised the lack of an empirically testable threshold for when a situation is self-relevant enough to qualify as a motivated performance situation in which CAT states may be analysed. The Blascovich group (2003) responded by citing one of their studies showing that goal relevance and task engagement can be easily manipulated (Blascovich et al., 1999). Precisely, the study compared performance in front of an audience with performance alone and observed that task engagement was greater when an audience was present, which they explained with evaluation apprehension concerns and associated increases in self-relevance. Blascovich and colleagues (2003) further stated that task engagement and goal relevance may be determined by multiple factors (e.g., financial, social, or personal motivators), a single one of which may already be sufficient to produce significant task engagement/goal relevance. Thus, the BPSM conceptualises task engagement and goal relevance as relative phenomena that depend on multiple situational factors, not all of which have been explored yet. However, it does provide an avenue for testing (i.e., by examining change in HR/VC reactivity) and manipulating (e.g., Blascovich et al., 1999) the self-relevance of a motivated performance situation.

Wright and Kirby (2003) also criticised the cardiovascular predictions of the BPSM. They pointed out that sympathetic-adrenomedullary activation (which the BPSM specifies to be involved in a challenge state) also includes the release of norepinephrine, which acts as a vasoconstrictor and therefore should oppose the vasodilation observed in a challenge state. To this apparent inconsistency, Blascovich and colleagues (2003) responded that epinephrine, which is considered to be the main vasodilator in a challenge state, acts to inhibit norepinephrine release (Brownley, Hurwitz, & Schneiderman, 2000). However, they acknowledged that other biological factors play a role in vasodilation and that the BPSM may therefore be overly simplistic regarding vasodilation and constriction in CAT states. Furthermore, Wright and Kirby criticised that the BPSM did not sufficiently explain blood pressure responses in motivated performance situations, to which Blascovich and colleagues responded that blood pressure variables are not definitive indicators of the cardiovascular CAT patterns, with blood pressure changes in both directions being possible (i.e., higher blood pressure in a challenge or a threat state). However, they acknowledged that early specifications of little to no blood pressure change in challenge states were likely too general.

Finally, Wright and Kirby (2003) argued against the BPSM's prediction that task engagement (which they, in an apparent misinterpretation of the model, referred to as an index of effort) will be independent of CAT states. The Blascovich group (2003) responded to this issue by rejecting the notion that HR increases could serve as a pure measure of invested effort and cited research in which participants invested similar effort, but exhibited large cardiovascular differences due to factors related to task engagement and CAT states (Mendes et al., 2002). In sum, the Blascovich group was not convinced that Wright and Kirby's critiques presented serious flaws of the BPSM and continued using it as a theoretical framework, whereas Wright and Kirby did not publish any more critiques of the BPSM.

1.5 Other Relevant Theories

Several other theoretical models incorporating the concepts of CAT states exist that differ from the BSPM in critical respects. For instance, the Theory of Challenge and Threat States in Athletes (TCTSA; Jones, Meijen, Sheffield, & McCarthy, 2009) also conceptualises CAT states as a single bipolar continuum that occurs in motivated performance situations. The TCTSA supports the predictions of the BPSM that performance, mental, and physical health should be better in a challenge than in a threat state. Furthermore, it predicts that CAT states influence attention, decision-making, invested effort, emotions, and interpretations of emotions. These predictions have been supported in several studies, indicating that better emotional and attentional functioning in a challenge, relative to a threat state. For example, after using instructional sets to manipulate participants into a challenge and a threat group, Moore and colleagues found that challenged participants, relatively to threatened participants, reported more favourable emotions, less anxiety, and more facilitative interpretations of anxiety (Moore et al., 2012; study 2, Moore, Wilson, et al., 2013).

However, the TCTSA is different from the BPSM in several ways. It exclusively predicts sport performance, which the BPSM only covers if the performance is non-metabolically demanding. It avoids discussing task engagement as a prerequisite to CAT states, instead specifying increased (versus decreased) task engagement as a consequence of a challenge (versus a threat) state. The TCTSA also takes a slightly different perspective on how antecedents relate to demand and resource evaluations, specifying separate respective antecedents for demand and resource appraisals. These antecedents include the perception and assessment of danger, uncertainty, and required effort for demand appraisals, as well as self-efficacy, perceived control, and achievement goal orientation for resource appraisals. Some of these antecedents were supported in subsequent studies, for example resource appraisals (Turner, Jones, Sheffield, Barker, & Coffee, 2014) or perceived required effort (Moore et al., 2014). It should be noted that although the TCTSA uses the word "appraisal" instead of "evaluations", it is consistent with the BPSM in its conceptualisation of CAT states as opposite ends to a single bipolar continuum.

Similarly, the integrative framework of stress, attention, and visuomotor performance (subsequently termed "the integrative framework") focuses on predicting the performance of visually guided motor skills (Vine, Moore, & Wilson, 2016). It employs a BPSM framework to include CAT states as predictors of attentional systems, which in turn predict the efficiency of visuomotor control, and visuomotor performance in motivated performance situations. Precisely, it states that a challenge state will be associated with a balanced reliance on the goal-directed (top-down) and the stimulus-driven (bottom-up) attentional systems (Corbetta & Shulman, 2002), whereas a threat state will be associated with heavier reliance on the stimulus-driven system. According to the integrative framework, relying too much on the stimulus-driven system will lead athletes to be distracted by irrelevant and/or threatening stimuli. A recent study supported this prediction by showing a cardiovascular response consistent with a challenge state was associated with longer quiet eye durations and lower search rates in a football penalty shooting task than a threat state, both of which are considered to reflect optimal goal-directed system use (Brimmell et al., 2019). However, the same study found conflicting results regarding the focus on distracting or threatening stimuli, as a challenge state was associated with more time spent fixating on the goalkeeper, who was considered a distracting stimulus.

The integrative framework also specifies that the performance outcome and the psychophysiological state will feed back into the demand and resource evaluation process, such that poor performance or a threat response will reduce resources relative to demands in future situations with comparable tasks. Conversely, a good performance outcome increases resources relative to demands in future situations with comparable tasks. The study by Brimmel and colleagues (2019) explicitly tested this prediction without finding support for it, albeit cognitive evaluations in the first penalty shooting trial predicted cognitive evaluations in the second trial. Furthermore, the integrative framework suggests that the negative effects of a threat state can be mitigated with compensatory strategies, for example by reappraising anxiety symptoms or by increasing effort. This prediction has not yet been explicitly tested, but it was based on previous research that found supportive results for this prediction, where arousal reappraisal instructions were successful in mitigating a threat state by producing cardiovascular reactivity more consistent with a challenge state than control instructions (Moore et al., 2015; Sammy et al., 2017). In sum, the integrative framework provides a perspective compatible with the BPSM, but several of its predictions remain to be supported by forthcoming empirical studies.

Unlike the previous two theories, Skinner and Brewer's (2004) adaptive approaches to competition model incorporates CAT states in a framework that is inconsistent with the BPSM. In particular, it uses Lazarus and Folkman's (1984) cognitive appraisal theoretical conceptualisation of CAT states (i.e., not as opposite ends to a continuum but as separate cognitive appraisals). However, Skinner and Brewer's (2004) model is consistent with the BPSM and the TCTSA in that it also finds challenge appraisal to relate to better performance (as BPSM) and positive emotion (as TCTSA), whereas threat appraisal relates to negative emotion (i.e., anxiety).

1.6 CAT Interventions

Several studies have examined interventions to directly or indirectly manipulate CAT states. To my knowledge, all studies except for two validation studies by Tomaka and colleagues (1997) have used a psychology-level approach to manipulating CAT states. The most direct approach to manipulating CAT states may have been to use instructional sets framing the task as a challenge or a threat, using cognitive- (e.g., Feinberg & Aiello, 2010; Drach-Zahavy & Erez, 2002) and cardiovascular-level (e.g., Turner et al., 2014; Tomaka et al., 1997) manipulation checks. In a typical example of this type of manipulation (Tomaka et al., 1997), participants were encouraged to think of the task as a challenge and think of themselves as someone capable of meeting that challenge to elicit a challenge state. To elicit a threat state, participants were informed that it was important to complete the task as quickly and accurately as possible, and that performance would be scored for speed and accuracy. Results showed that these instructional sets were generally effective in eliciting both cognitive evaluations and cardiovascular responses consistent with CAT states.

Studies have also manipulated CAT states with established interventions or techniques that do not explicitly focus on CAT states. For example, a study by Moore and colleagues found that quiet eye training elicited cognitive evaluations more consistent with a challenge state (by increasing resource evaluations) and improved performance (Moore, Vine, Freeman, & Wilson, 2013). Three studies have successfully used arousal reappraisal instructions to elicit cardiovascular responses (Jamieson, Nock, & Mendes, 2012; Moore et al., 2015; Sammy et al., 2017) and resource evaluations (Sammy et al., 2017) reflective of a relative challenge state (compared to control instructions). Another study found that self-distancing produced cardiovascular responses consistent with relative challenge during a speech about personal qualifications for one's dream job (Streamer, Seery, Kondrak, Lamarche, & Saltsman, 2017).

Furthermore, social manipulations have also been found to affect cardiovascular CAT responses, as experiments on social facilitation, social feedback, solo status, group status, rejection, stereotype threat, and stigma have shown (Blascovich, Mendes, Hunter, Lickel, & Kowai-Bell, 2001; Blascovich, Mendes, Hunter, & Salomon, 1999; Kassam, Koslov, & Mendes, 2009; Mendes, Blascovich, Hunter, Lickel, & Jost, 2007; Mendes et al., 2002; Mendes, McCoy, Major, & Blascovich, 2008; Scheepers, 2017; White, 2008). For example, Mendes and colleagues (2002) found that interacting with a Black confederate or a confederate from a socioeconomically disadvantaged group provoked a cardiovascular threat response in non-Black participants, whereas participants exhibited a cardiovascular challenge response when interacting with a White confederate. Blascovich and colleagues (1999) found that a cardiovascular challenge response could be elicited by asking participants to perform a well-learned task in front of an audience, whereas a threat response could be elicited by asking participants to perform an unlearned task in front of an audience. A last example from this group of studies showed that nonverbal feedback from interviewers in a mock job interview was able to manipulate cardiovascular CAT responses, with positive nonverbal feedback (e.g., nodding, smiling) eliciting a relative challenge and negative nonverbal feedback (e.g., head-shaking, arm-crossing) eliciting a relative threat cardiovascular response (Kassam et al., 2009).

1.7 Future Research Directions

The literature indicates that the BPSM of CAT has received substantial empirical support as a theory. However, there are some areas that it has not yet elucidated. For example, research on the variability and dynamic nature of CAT states is still scarce. Although the terminology indicates that there may be a lively dynamic to CAT *states*, few studies have followed a group of individuals and measured CAT states at multiple time points. Given the specification of the BPSM that motivated performance situations are subject to continuous reappraisal as new information is obtained (Blascovich, 2008a), one should expect to find some within-subjects variation in CAT states. Indeed, some research has found that CAT states measured before performance were associated with CAT evaluations after performance. One study found that pre-speech task evaluations and objectively rated speech performance predicted post-speech task evaluations (Rith-Najarian, McLaughlin, Sheridan, & Nock, 2014). Another study instructed participants to complete four mental arithmetic tasks and measured cardiovascular CAT variables throughout (Quigley, Barrett, & Weinstein, 2002). The results showed that task-related behaviours and cardiovascular reactivity predicted post-task evaluations. These findings suggest that reappraisal may occur as a function of cognitive evaluations before, as well as cardiovascular responses and behavioural outcomes during the task, although the correlational nature of the studies prevented any definitive inference of causality. Research also showed that different tasks may provoke differential relationships between cognitive and cardiovascular indicators of CAT states, thereby offering a varying degree of support for the predictions of the BPSM (Trotman, Williams, Quinton, & Veldhuijzen van Zanten, 2018). Precisely, it was found that the relationships between cognitive CAT appraisals (pre-task & post-task) and cardiovascular CAT responses were more consistent with the BPSM on a computer-based competitive task, relative to a public speaking task. However, this finding is limited by the researchers' conceptual adherence to cognitive appraisal theory (Lazarus & Folkman, 1984), as they did not use a cognitive evaluation score (e.g., Brimmel et al., 2018), but independent CAT appraisals.

The three previously mentioned studies did not measure CAT states on multiple days, although time should also be considered when examining CAT states in motivated performance situations. One seminal study measured CAT states at the beginning of the season in student softball and baseball athletes and was able to predict average season performance, with a challenge state at the beginning of a season relating to better average performance than a threat state (Blascovich et al., 2004). Potentially inspired by this finding, Tomaka and colleagues recently proposed a questionnaire to measure individual differences in CAT evaluations; that is, a general tendency to respond to a given motivated performance situation with a challenge or a threat state (Tomaka, Palacios, Champion, & Monks, 2018). The questionnaire comprises six dimensions (conflict situations, unexpected events, public speaking, transport, social anxiety, and financial concerns) with four items each, thus containing 24 items in total. The questionnaire exhibited high internal consistency, predicted depression and post-traumatic stress disorder symptoms above the Perceived Stress Scale, and exhibited a negative association with life satisfaction (indicating that a tendency toward threat relates to lower life satisfaction).

However, despite the first evidence for stable CAT tendencies, one should also expect some variation over longer time periods due to the many entropic influences in the pursuit of big overarching goals that comprise many constituent motivated performance situations. For example, the Brazilian national team players in the 2014 football world championships might have experienced relative challenge in the early stages of the tournament until the unexpected injury of their eminent player Neymar Jr. provoked a sudden threat state. Hence, future research should examine CAT states across different motivated performance situations (i.e., different tasks and/or time points) to explore the interplay of stable CAT dispositions and situational factors (i.e., between-subjects versus within-subjects variation). For example, research on police officers' stress appraisals showed that the stressor and the individual may interact in explaining stress appraisals (Lucas, Weidner, & Janisse, 2012). In particular, this research found that there were some general differences between officers, some general differences between stressful scenarios, but also some differences between some officers/scenarios in some, but not other scenarios/officers regarding how they appraised stressors. Hence, there should be no either-or question, but rather the question of how individual difference and situational factors interact to explain CAT states and their effects on performance, health, and other outcomes.

More research could examine ways of experimentally manipulating CAT states. In particular, future intervention studies could explore whether there are alternative or complementary intervention approaches apart from the commonly researched psychology-level approach. For example, since the BPSM specifies a catecholamine involvement in a challenge state, research could test whether catecholamine agonists or precursors may influence CAT states. Also, intervention research could test established sport psychological interventions (e.g., goal-setting, self-talk) that have not yet been examined regarding their potential to promote influence CAT states.

1.8 Conclusion

The extant literature describing and testing the BPSM of CAT spans over two decades and suggests that the model is useful in various theoretical and applied settings, including the discipline of sport and exercise psychology. It has withstood theoretical criticism and empirical scrutiny, but some areas of the model remain understudied. Hence, future research could focus on the dynamic nature of CAT states and examine trait and state components of CAT states in different persons, on different tasks, and at different time points. Furthermore, more research could test CAT interventions to explore alternatives to previously tested psychology-level CAT interventions. This thesis aims to contribute to the knowledge base by addressing some of the gaps in the literature identified by this chapter. To do this, it reports on a systematic literature review and four empirical studies that examined novel research questions in the field of CAT research using a BPSM framework. The first two empirical studies examined the generalisability of CAT states across tasks and time points, and the last two empirical studies examined the previously untested impact of interventions on CAT states. Table 1.1 lists rationales and aims for each of the upcoming thesis chapters. The next chapter reports on the systematic review of the literature describing the relationship between CAT states and performance.

Table 1.1

Chapter Rationales, Aims, and Methods

Chapter	Rationale	Aim	Method
2	Several studies have examined the relationship between CAT states and performance, but no publication has reported a systematic review of this relationship.	To systematically review the relationship between CAT states and performance in the published literature.	Systematic literature review
3	No previous study has decomposed the variance in CAT states across repeated measurements.	To partition the variance in CAT states into personal (person), situational (task, week), and interaction components.	Observational study involving 12 repeated measurements (three weeks, four tasks)
4	Results of chapter 3 were obtained in a laboratory-based context; no previous study has examined whether findings generalise to an elite sport context.	To partition the variance in CAT states into personal (athlete), situational [competition(athlete)], and interaction components.	Observational study involving three repeated measurements at three out of six competitions in a nested design
5	The effects of instructional and motivational self-talk partially overlap with those of a challenge state. This overlap might be explained by an ability of instructional and motivational self-talk to promote a challenge state.	To examine whether instructional and motivational self-talk promote a challenge state.	Experiment involving two measurements (three-group between-subjects design)
6	The BPSM specifies a catecholamine involvement in a challenge state. The catecholamine precursor tyrosine has been found to raise serum tyrosine and catecholamine levels. Positive effects of tyrosine on cognitive and motor performance might therefore be due to tyrosine promoting a challenge state.	To examine whether tyrosine intake promotes a challenge state.	Experiment involving two measurements (two-condition within-subjects design)
7	The BPSM specifies that cognitive and cardiovascular CAT variables are positively associated. The present studies have collected data to test this prediction, but not explicitly tested it individually. Thus, a meta-analysis could provide robust evidence to test the prediction.	To conduct a meta-analysis of the relationship between cognitive and cardiovascular indicators of CAT states collected in all empirical studies of this thesis.	Meta-analysis

Chapter 2

The Relationship between CAT States and Performance: A Systematic Review

Published as:

Hase, A., O'Brien, J., Moore, L. J., & Freeman, P. (2019). The relationship between challenge and threat states and performance: A systematic review. *Sport, Exercise, and Performance Psychology*, *8*, 123-144. doi:10.1037/spy0000132

2.1 Abstract

The biopsychosocial model of challenge and threat states specifies that these states engender different physiological and behavioural responses in potentially stressful situations. This model has received growing interest in the sport and performance psychology literature. The present systematic review examined whether a challenge state is associated with superior performance than a threat state. Across 38 published studies that conceptualised challenge and threat states in a manner congruent with the biopsychosocial model, support emerged for the performance benefits of a challenge state. There was, however, significant variation in the reviewed studies in terms of the measures of challenge and threat states, tasks, and research designs. The benefits of a challenge state on performance were largely consistent across studies using cognitive, physiological, and dichotomous challenge and threat measures, cognitive and behavioural tasks, and direct experimental, indirect experimental, correlational, and quasi-experimental designs. The results imply that sports coaches, company directors, and teachers might benefit from trying to promote a challenge state in their athletes, employees, and students, respectively. Future research could benefit from a greater consensus on how best to measure challenge and threat states to help synthesise the evidence across studies. Specifically, we recommend that researchers use both cognitive and physiological measures and develop stronger manipulations for experimental studies. Finally, future research should report sufficient information to enable risk of bias assessment.

2.2 Introduction

Understanding individuals' responses to stress is key for optimising performance in contexts including business, medicine, education, and sport. Although some models explain individuals' successes and failures in terms of psychology or physiology, one increasingly popular theory combines these perspectives. The BPSM of CAT states (Blascovich & Mendes, 2000) built on Lazarus and Folkman's (1984) transactional theory of stress and Dienstbier's (1989) theory of physiological toughness, and has been applied to contexts as diverse as sport, education, and medicine (Moore, Wilson, et al., 2013; Roberts, Gale, McGrath, & Wilson, 2015; Seery, Weisbuch, Hetenyi, & Blascovich, 2010). Across these contexts, CAT states have been associated with different performance outcomes (e.g., Allen & Blascovich, 1994; Blascovich et al., 2004), although some studies have found non-significant or contradictory results (e.g., Feinberg & Aiello, 2010; Laborde, Lautenbach, & Allen, 2015), and there is notable diversity in how CAT states have been measured and the research designs employed. To advance our understanding of the impact of CAT states on performance, the consistency of findings across different methods, and to highlight important directions for future research, the current article reports a systematic review of the published literature that utilised the BPSM as a theoretical framework.

Central to the BPSM is the assumption that CAT states only occur in motivated performance situations. Motivated performance situations are goal-relevant, evaluative, and potentially stressful, requiring adequate active performance in order to ensure wellbeing and personal growth (Blascovich & Mendes, 2000). Sport competitions, academic exams, and job interviews are typical examples of such situations. Importantly, according to the BPSM, CAT states represent opposite ends of a unidimensional continuum rather than two dichotomous states, allowing researchers to examine relative (rather than absolute) differences in challenge and threat (i.e., greater vs. lesser challenge or threat; Blascovich, 2008a). This contrasts the earlier views of Lazarus and Folkman (1984), and other researchers (e.g., Skinner & Brewer, 2004), who considered CAT as independent cognitive appraisals that can occur simultaneously. Although these other frameworks offer useful insights, this review focused only on publications that examined CAT states in the unidimensional manner hypothesised in the BPSM.

CAT states differ in terms of underlying cognitive evaluations and resulting physiological responses, which are predicted to be linked (Blascovich & Mendes, 2000). According to the BPSM, challenge states are characterised by the largely subconscious evaluation that one's personal coping resources match or exceed situational demands. Physiologically, challenge states are marked by increases in HR and CO, and decreases TPR. This cardiovascular pattern is due to sympathetic-adrenomedullary activation, which causes epinephrine release, and dilation of the blood vessels. In contrast, threat states are characterised by an evaluation that coping resources fall short of situational demands. Threat states are indexed by little change or small increases in HR, little change or minor decreases in CO, and little change or small increases in TPR (Tomaka et al., 1993). This physiological response is due to additional activation of the pituitary-adrenocortical pathway, which constricts blood vessels, causes cortisol release, and inhibits the effects of sympathetic-adrenomedullary activation (Blascovich & Mendes, 2000). Importantly, validation studies showed that: a) cognitive CAT evaluations and physiological CAT responses were significantly correlated, and b) cognitive CAT evaluations triggered physiological responses, not vice versa (Blascovich, 2008a). These divergent CAT states are predicted to influence performance, with challenge states being related to superior performance than threat states.

The relevance of the BPSM to a range of contexts has led to considerable variation in the tasks and performance outcomes examined across the literature. For example, studies have examined the relationship between CAT states and cognitive performance in academic (Seery et al., 2010), GRE word problem (Chalabaev, Major, Cury, & Sarrazin, 2009), and mental arithmetic (Kelsey et al., 2000) tasks. Further, Blascovich et al. (2004) found that a cardiovascular CAT index, measured during a pre-season speech about athletes' sports, predicted batting performance during the season, with a challenge state linked to better performance than a threat state (i.e., more runs). This initial evidence provided impetus for subsequent research involving behavioural tasks as varied as simulated surgery (Vine et al., 2013) and cricket batting (Turner et al., 2013).

This early research also led to the development of new theories that extended the predictions of the BPSM (i.e., Theory of Challenge and Threat States in Athletes [TCTSA]; Jones et al., 2009; integrated framework of stress, attention, and visuomotor performance; Vine et al., 2016). These theories suggest that CAT states could influence performance through various mechanisms. For example, the TCTSA predicts that a threat state may lead to more negative emotions, unfavourable interpretations of emotions, impaired cognitive functioning, decision-making and anaerobic power, greater self-regulation, increased reinvestment and avoidance coping, and less effective attention, which may in turn impair performance (Jones et al., 2009). Further, Vine et al. (2016) argue that a threat state might deter performance by disrupting attentional and visuomotor control, causing individuals to become distracted by less relevant (and potentially negative) stimuli at the expense of more important task-relevant cues. This is in keeping with the original mechanism proposed by Blascovich et al. (2004), who speculated that attentional resources might be diverted from the task at hand towards the environment or themselves during a threat state. However, to date, relatively little research has tested these potential mechanisms (e.g., Moore et al., 2012).

With increasing interest in the BPSM, there has been greater diversity in the conceptualisation and measurement of CAT states. Indeed, while some authors have used self-report measures of demand and resource evaluations (e.g., Gildea, Schneider, & Shebilske, 2007), others have used physiological indices computed from CO and TPR reactivity (i.e., change in CO and TPR from baseline to post-instruction/task exposure; e.g., Blascovich et al., 2004). Although both the cognitive evaluations and physiological responses accompanying CAT states are predicted to influence performance, it is not known which has the strongest effect. Even within these approaches, little consensus exists regarding standardised measurements. For example, both single- and multi-item self-report measures of cognitive evaluations have been used to calculate either a ratio (e.g., demands divided by resources), or a difference score (e.g., resources minus demands). Researchers have also differed in the timing and duration of baseline and post-instruction/task exposure periods when recording cardiovascular data, and have used different methods to calculate a single CAT index from CO and TPR reactivity (e.g., difference vs. residualised change scores).

In addition to the diversity in the measurement of CAT states and the tasks employed, studies have adopted different research designs. Some studies have employed experimental designs, directly manipulating individuals into CAT states and observing performance. For example, Moore and colleagues used verbal instructions to elicit CAT states before a golf putting task, and found that the golfers in the challenge group outperformed those in the threat group (Moore, Wilson et al., 2013). Other experimental studies have indirectly manipulated CAT states via an antecedent and then measured performance (e.g., resource appraisals; Turner et al., 2014). Correlational studies have also been employed, with CAT states observed before a task and subsequently related to performance (e.g., Turner et al., 2013). Finally, studies have used quasi-experimental designs, recording CAT states with continuous measures, and then splitting the sample into CAT groups before examining between-group differences in performance (e.g., via median split; Gildea et al., 2007).

Given the increasing adoption of the BPSM for understanding performance variation during stressful tasks, aligned with notable diversity in the conceptualisation of CAT states, performance outcomes, and research designs employed, the primary aim of this systematic review was to examine the pattern of associations between CAT states and performance outcomes. The secondary aim was to examine the consistency of this pattern across different conceptualisations of CAT states (i.e., cognitive evaluations vs. physiological responses vs. dichotomous groups), performance outcomes (i.e., cognitive vs. behavioural tasks), and research designs (i.e., direct experimental vs. indirect experimental vs. correlational vs. quasi-experimental designs). Synthesising the current evidence will provide crucial insight into the utility of the BPSM to explain performance variation under stress, the impact of employing different methods, and highlight important directions and methodological considerations for future research.

2.3 Method

This systematic review was conducted in accordance with the Preferred Reporting Items for Systematic reviews and Meta-Analyses guidelines (Moher, Liberati, Tetzlaff, &

Altman, 2009). It involved four steps: (1) initial literature search (including selection of search terms, electronic databases, and inclusion criteria), (2) screening based on title, (3) screening based on abstract, and (4) screening based on full text. Two independent assessors completed each step, compared their records and discussed any disagreements. The assessors searched for relevant articles using the following databases: MedLine, PsycINFO, and SPORTDiscus (combined in one search) and Web of Science (in a separate search). The search terms were ("challenge and threat" AND "performance"). To be included, studies had to fulfil five inclusion criteria: (1) published in English in a peer-reviewed academic journal, (2) report at least one empirical study, (3) conducted with healthy human participants, (4) conceptualise CAT in terms of a unidimensional continuum, and (5) report at least one performance outcome and its association with at least one CAT measure, or dichotomous CAT groups that were compared on a CAT measure in a manipulation check.

To examine the consistency of the pattern of associations between CAT states and performance within different conceptualisations of CAT states, performance outcomes and research designs, we used Sallis, Prochaska, and Taylor's (2000) sum code classification. This classification focuses on the percentage of studies that demonstrate a statistically significant effect. Further, to assess the quality and risk of bias in experimental and non-experimental studies, respectively, the Cochrane Collaboration's tool for assessing risk of bias (Higgins & Altman, 2008) and the Risk of Bias Assessment Tool for Nonrandomised Studies (Kim et al., 2013) were used. For experimental studies, two independent assessors examined random sequence generation (were experimental conditions assigned randomly?), allocation concealment (could condition allocations have been foreseen before/during enrolment?), blinding of participants and personnel (were participants and researchers blind to the participants' allocated experimental condition?), blinding of outcome assessment (were outcome assessors blind to experimental condition?), incomplete outcome data (were attrition/exclusion rates and reasons reported?), selective reporting (was there a possibility of selective reporting?), and other sources of bias (Higgins & Altman, 2008). For non-experimental studies, two independent assessors examined blinding of outcome assessment, incomplete outcome data, selective reporting, selection of participants (how adequate was the selection of participants?), confounding variables (was there adequate consideration of confounders?), and intervention (exposure) measurement (was there performance bias caused by inadequate measurement of exposure?; Kim et al., 2013).

2.4 Results

The initial search (conducted in December 2017) yielded 1107 unique results. After reviewing titles, 155 records remained. After reading abstracts, 59 records remained. After reviewing full-texts, 30 articles reporting 38 studies with a total of 3257 participants were identified and included in the review. Figure 2.1 illustrates the search and screening process. Inter-rater agreements in the second, third, and fourth step were 96.6%, 84.4%, and 84.7%. Disagreements were resolved through discussion between the assessors and a third member of the research team.

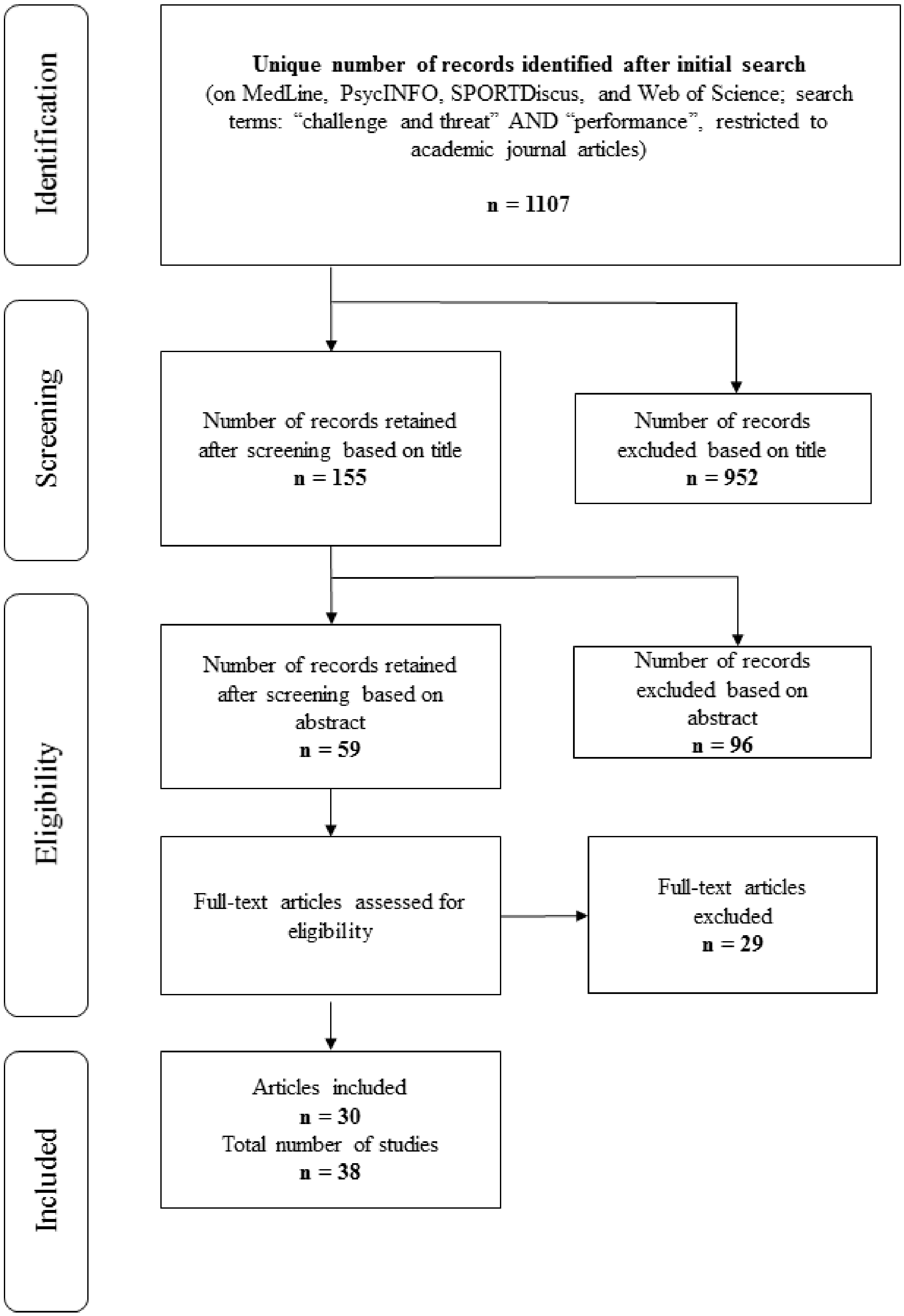

Figure 2.1. Systematic review search and screening procedure.

2.4.1 General Study Characteristics

Table 2.1 presents the characteristics and main outcomes of the included studies. Sample sizes ranged from 16 to 238 with a mean sample size of 85.7 participants (*SD* = 54.4). Most samples contained both genders, but four samples were all male (Gildea et al., 2007; Laborde et al., 2015; Turner et al., 2013), and five samples were all female (Chalabaev et al., 2009; Chalabaev, Major, Sarrazin, & Cury, 2012; Mendes et al., 2007; Study 2, Scheepers, 2017; Turner, Jones, Sheffield, & Cross, 2012). The average age in the 28 studies that reported this statistic ranged from 11.0 to 36.3 years with an average mean of 22.5 years (*SD* = 4.9). The remaining studies reported a mode age of 18 years (Quigley et al, 2002), a median of 28 years (Roberts et al., 2015), or no age statistic (Blascovich et al., 2004; Chalabaev et al., 2009; Chalabaev et al., 2012; Feinberg & Aiello, 2010; Kelsey et al., 2000; Seery et al., 2010). Most studies sampled university students, but others incorporated athletes, doctors, adolescents, academic staff, and non-specified adults.

Table 2.1

Summary of Included Studies

Reference Number	Authors, Year	N	Design	Population	Mean age (years)	CAT	Main Performance Measures	Results	Effect Sizes
1	Blascovich, Seery, Mugridge, Norris & Weisbuch, 2004	27	CR	Baseball and softball student athletes	N/A	P	Baseball and softball season performance (runs created)	CAT index related to runs created during season; (challenge > threat)	R^2 =.11
2	Chalabaev, Major, Cury & Sarrazin, 2009	27	EX - performance goal	Female undergraduates	N/A	P, C	Multiple-choice score on GRE word problems	Self-reported challenge was unrelated to performance CO and TPR were related to performance, but only examined separately (no CAT index)	N/A
3	Chalabaev, Major, Sarrazin & Cury, 2012	58	EX - Performance goal (approach, avoidance, control)	Female psychology undergraduates	N/A	C	Score on math word problems from GRE practice book	For those participants who received a performance avoidance goal, challenge was associated with better performance than threat	R^2 =.06
4	Feinberg & Aiello, 2010[1]	91	EX - CAT appraisal	Undergraduates	N/A	C, DC	Mental arithmetic score	Threat group outperformed challenge group	*d* = 0.85

[1] Studies 1, 2, and 4 from this publication were included in the systematic review. Study 3 was not included because it did not report the results of the

		238	EX - CAT appraisal		N/A	C, DC	Mental arithmetic score Anagram task score	Challenge group outperformed threat group No significant difference between groups	N/A
		54	EX - CAT appraisal		N/A	C, DC			N/A
5	Gildea, Schneider & Shebilske, 2007	54 154 48	QE QE QE	Adults and adolescents (all male in studies 1 and 3)	22.5 19.9 24.1	C, DC C, DC C, DC	Space Fortress (total scores; used in all studies)	Challenge associated with higher scores than threat across three experiments (not significant in experiment 2)	*d* = 1.09 *d* = 0.29 *d* = 0.65
6	Kelsey et al., 2000	162	CR	Psychology undergraduates	N/A	C	Three arithmetic tasks (number of responses, arithmetic errors)	Number of responses inversely correlated with pre-task evaluations (challenge > threat) Arithmetic errors positively correlated with pre-task evaluations	N/A N/A
7	Laborde, Lautenbach & Allen, 2015	96	CR	Male sport science students	24.8	C	Concentration grid exercise (consecutive numbers clicked in two minutes)	CAT not significantly related to visual search task performance	N/A

---

main effect comparison between the CAT conditions.

8	Mendes, Blascovich, Hunter, Lickel & Jost, 2007	47	EX - 2x2 (confederate ethnicity x confederate accent)	Female students	19.6	P	Word-finding task (number and accuracy of responses)	No significant effect of CAT index on performance in a mediation model (marginally significant trend was found)	N/A
9	Moore, Vine, Freeman & Wilson, 2013	30	EX - training (quiet eye, technical)	Undergraduates without golf putting experience	19.7	C	Golf putting (mean radial error)	Evaluations mediated the relationship between group and mean radial error (challenge associated with smaller radial error than threat)	N/A
10	Moore, Vine, Wilson & Freeman, 2012	127	EX – CAT appraisal	Undergraduates without golf putting experience	19.5	P, C, DC	Golf putting (mean radial error)	Lower mean radial error in challenge group	*d* = 0.69
11	Moore, Vine, Wilson & Freeman, 2014	120	EX - 2x2 (effort x support)	Undergraduates	21.6	P, C, DC	Laparoscopic surgery completion time	Low effort group (challenged) outperformed high effort group (threatened)	η^2_p = .12
12	Moore, Vine, Wilson & Freeman, 2015	50	EX - Arousal reappraisal	Participants without golf putting experience	20.2	P, DC	Golf putting (mean radial error)	Arousal reappraisal group was more challenged and performed more accurately (lower error)	*d* = 0.93
13	Moore, Wilson, Vine, Coussens & Freeman, 2013	199	CR	Competitive golfers	36.3	C	Golf competition performance	Challenge evaluations were associated with superior competition	R^2 = .09

		60	EX – CAT appraisal	Experienced golfers	22.9	P, C, DC	Golf putting (putts holed, performance error)	performance than threat evaluations Challenge group holed higher percentage of putts than threat group Challenge group had lower error than threat group	*d* = 0.63 *d* = 0.70
14	Moore, Young, Freeman & Sarkar, 2018	100	CR	Participants engaging in club or university level sports	21.9	P, C	Dart-throwing task	Physiological CAT index and cognitive CAT evaluations related to dart-throwing performance (challenge > threat)	R^2 = 0.08 R^2 = 0.11
15	O'Connor, Arnold & Maurizio, 2010	138	EX - academic focus	Undergraduates	24.8	C	Negotiation task score	Threat associated with lower negotiation outcomes than challenge	R^2 = .16
		196	EX - 2x2 (CAT appraisal x task structure)	Undergraduates	22.2	C, DC	Negotiation task score	Challenge group scored better negotiation outcome than threat group in the integrative task structure condition only – no main effect	*d* = 0.32
16	Quigley, Barrett & Weinstein, 2002	74	CR	Psychology undergraduates	18 (mode)	P, C	Four verbal mental arithmetic tasks (attempts, number correct)	No relation between cognitive evaluations and performance (number of attempts made, percentage correct responses) No analysis reported for physiological data	N/A

17	Rith-Najarian, McLaughlin, Sheridan & Nock, 2014	79	CR	Adolescents	14.70	P, C	Independently rated speech performance	No relation between physiological and cognitive measures of CAT and performance before task	N/A
18	Roberts, Gale, McGrath & Wilson, 2015	94	CR	Doctors	28 (median)	C	Overall station performance score	CAT predicted station performance (threat < challenge)	N/A
19	Sammy et al., 2017	54	EX – Arousal reappraisal	Undergraduates	21.7	P, C, DC	Dart-throwing task	Arousal reappraisal group more challenged on physiological index and evaluations, but not better on dart-throwing task	N/A
20	Scheepers, 2017	103	EX – 2x2 (Group status x group legitimacy)	Female undergraduates	21	P, DC	Pattern recognition task	CAT index negatively correlated with performance (higher challenge – lower response times)	R^2 = 0.07
								High status group was more challenged and outperformed low status group	N/A
21	Schneider, 2004	59	QE	Undergraduates	21	C, DC	Mental arithmetic performance (responses, errors)	Threat group gave fewer responses	d = -0.78
								Threat group made more errors	d = 0.53
								CAT predicted percent correct (threat < challenge)	r = -.33

22	Schneider, Rench, Lyons & Riffle, 2012	152	CR	Psychology undergraduates	20.3	C	Mental arithmetic score (responses and accuracy)	Cognitive evaluations were negatively related with performance (threat < challenge)	N/A
23	Scholl, Moeller, Scheepers, Nuerk & Sassenberg, 2017	50	CR	Undergraduates	20.0	P	Number bisection task[2] errors made	Physiological CAT index was negatively related with number of errors made in all task conditions (challenge associated with less errors than threat)	R^2 = .21 R^2 = .20 R^2 = .11 R^2 = .16
24	Seery, Weisbuch, Hetenyi & Blascovich, 2010	95	CR	Undergraduates	N/A	P	University course grades	Cardiovascular CAT (academic interests speech) predicted course grades (challenge > threat) No association found for general test taking speech	sr^2 = .04 N/A
25	Turner, Jones, Sheffield, Barker & Coffee, 2014	46	EX - resource appraisals	Undergraduates and academic staff	21.7	P, DC	Bean-bag throwing score	Performance not significantly higher in challenge group	d = 0.50
26	Turner, Jones, Sheffield & Cross, 2012	25	CR	Academic staff members	34.0	P, C	Modified Stroop accuracy and latency	Cardiovascular challenge responses predicted superior performance over threat responses in both studies	R^2 = .16
		21	CR	Female netball players	21.1	P, C	Netball shooting score		R^2 = .14

[2] Analyses were only provided for each of the four sub-conditions of the number bisection task. The authors did not report on a total performance score. Thus, four values are reported in the "Effect Sizes" column.

27	Turner et al., 2013	42	CR	Male elite-level cricketers	16.5	P, C	Cricket batting task (runs awarded by coaching staff)	Physiological CAT associated with batting performance (challenge > threat) Cognitive evaluations not associated with performance	N/A N/A
28	Vine, Freeman, Moore, Chandra-Ramanan & Wilson, 2013	52	CR	Final-year medical students	20.5	P, C	Laparoscopic surgery task completion time	Cognitive evaluations associated with performance under pressure (challenge > threat) Relationship not mediated by physiological CAT index	N/A N/A
29	Vine et al., 2015	16	CR	Active pilots	34.8	C	Flight simulator metrics	Challenge evaluation associated with better performance than threat	R^2 = .61
30	White, 2008	128	EX - Solo status manipulation	Undergraduates	19.1	C	Math test scores Recall task score	Challenge associated with higher math test scores than threat Challenge was only associated with better performance than threat under solo status.	N/A N/A
		90	EX - Solo status manipulation		19.5	C	Math test score	Challenge associated with higher math test scores than threat	N/A

Note. CAT = Challenge and threat variables recorded. CR = Correlational. DC = Dichotomous (challenge group vs. threat group). EX = Experimental. QE = Quasi-experimental. C = Cognitive. P = Physiological.

2.4.2 Risk of Bias in Individual Studies

Table 2.2 presents the risk of bias results. Interrater agreements were 84.1% and 85.8% for experimental and non-experimental studies, respectively. The assessors resolved disagreements in discussions with a third member of the research team. In experimental studies, the lowest risk of bias ratings emerged for “random sequence generation”, “incomplete outcome data”, and “other sources of bias”, as 88.9%, 77.8%, and 100% of studies received a “low risk of bias” rating, respectively. Unclear risk of bias was more apparent for “allocation concealment”, “blinding of participants and personnel”, “blinding of outcome assessment”, and “selective reporting”, with 88.9%, 88.9%, 55.6%, and 100% of studies rated as “unclear risk of bias” respectively. The assessors rated one study (5.6%) in the “incomplete outcome data” category as “high risk of bias”.

In non-experimental studies, a low risk of bias ratings emerged for “blinding of outcome assessment”, “incomplete outcome data”, “confounding variables”, and “intervention (exposure) measurement”, as 55.0%, 75.0%, 100%, and 100% of studies in these categories received a “low risk of bias” rating, respectively. “Selective reporting” and “selection of participants” received mostly “unclear risk of bias” ratings (100% and 90.0%, respectively). The assessors rated two studies (10.0%) in the “incomplete outcome data” category as “high risk of bias”.

Table 2.2

Risk of Bias Assessment Results

Experimental Studies								
Reference Number		Random Sequence Generation	Allocation Concealment	Blinding of Participants and Personnel	Blinding of Outcome Assessment	Incomplete Outcome Data	Selective Reporting	Other Sources of Bias
2		Low	Unclear	Unclear	Unclear	Low	Unclear	Low
3		Low	Unclear	Unclear	Unclear	Low	Unclear	Low
4	Study 1	Low	Unclear	Unclear	Low	Low	Unclear	Low
	Study 2	Low	Unclear	Unclear	Low	Low	Unclear	Low
	Study 3	Unclear	Unclear	Unclear	Low	Low	Unclear	Low
8		Low	Low	Low	Unclear	Unclear	Unclear	Low
9		Low	Unclear	Unclear	Unclear	Low	Unclear	Low
10		Low	Unclear	Unclear	Unclear	Low	Unclear	Low
11		Low	Unclear	Unclear	Unclear	Low	Unclear	Low
12		Low	Unclear	Unclear	Unclear	Unclear	Unclear	Low
13	Study 2	Low	Unclear	Unclear	Unclear	Low	Unclear	Low
15	Study 1	Low	Unclear	Unclear	Low	Low	Unclear	Low
	Study 2	Low	Unclear	Unclear	Low	Low	Unclear	Low
19		Low	Unclear	Unclear	Unclear	Low	Unclear	Low
20		Low	Low	Low	Low	Unclear	Unclear	Low
25		Low	Unclear	Unclear	Unclear	High	Unclear	Low
30	Study 1	Unclear	Unclear	Unclear	Low	Low	Unclear	Low
	Study 2	Low	Unclear	Unclear	Low	Low	Unclear	Low
Non-experimental Studies								
		Blinding of Outcome Assessment	Incomplete Outcome Data	Selective Reporting	Selection of Participants	Confounding Variables	Intervention (Exposure) Measurement	
1		Low	Unclear	Unclear	Unclear	Low	Low	
5	Study 1	Low	Low	Unclear	Unclear	Low	Low	
	Study 2	Low	Low	Unclear	Unclear	Low	Low	
	Study 3	Low	Low	Unclear	Unclear	Low	Low	
6		Unclear	Low	Unclear	Unclear	Low	Low	
7		Low	Low	Unclear	Unclear	Low	Low	
13	Study 1	Low	Low	Unclear	Unclear	Low	Low	
14		Low	Low	Unclear	Unclear	L ow	Low	
16		Unclear	Low	Unclear	Unclear	Low	Low	
17		Unclear	Low	Unclear	Unclear	Low	Low	
18		Low	Low	Unclear	Low	Low	Low	
21		Unclear	High	Unclear	Unclear	Low	Low	
22		Unclear	Low	Unclear	Unclear	Low	Low	
23		Low	Unclear	Unclear	Unclear	Low	Low	
24		Unclear	High	Unclear	Unclear	Low	Low	
26	Study 1	Low	Low	Unclear	Unclear	Low	Low	
	Study 2	Unclear	Low	Unclear	Unclear	Low	Low	
27		Unclear	Unclear	Unclear	Unclear	Low	Low	
28		Unclear	Low	Unclear	Unclear	Low	Low	
29		Low	Low	Unclear	Low	Low	Low	

Note. For the "Reference Number" column coding, please consult the corresponding column in Table 2.1.

2.4.3 Association between CAT States and Performance

Of the 38 included studies, 28 (74%) found an effect on performance favouring a challenge state, although three of the observed effects were contingent on an interaction with another variable. The three interaction effects depended on solo status (performing alone or not; Study 1, White, 2008), performance goals (performance-avoidance or approach goal; Chalabaev et al., 2012), and integrative task structure (whether concessions on less important aspects of a negotiation tasks led to gains on more important aspects or not; Study 2, O'Connor, Arnold, & Maurizio, 2010). Of the remaining 10 studies, one found an effect favouring a threat state (Study 1, Feinberg & Aiello, 2010), and nine found no significant effects (Chalabaev et al., 2009; Study 4, Feinberg & Aiello, 2010; Study 2, Gildea et al., 2007; Laborde et al., 2015; Mendes et al., 2007; Quigley et al., 2002; Rith-Najarian et al., 2014; Sammy et al., 2017; Turner et al., 2014). At least one effect size was reported in 24 studies, yielding 29 in total: 12 Cohen's *d* values ranging from 0.29 to 1.09, 15 R^2 values ranging from .06 to .61, one sr^2 of .04, and one η_p^2 of .12 (see Table 2.1). These reflected 11 small, 14 medium, and four large effect sizes (Cohen, 1992).

2.4.3.1 Effects of cognitive, physiological, and dichotomous CAT measures on performance. Table 2.3 lists the associations between CAT states and performance based on whether CAT was analysed as a continuous cognitive, continuous physiological, or dichotomous variable. The dichotomous category included studies that compared challenge and threat groups in the analysis, regardless of whether the groups were created by an experimental manipulation or by a median split of a continuous CAT measure. Studies that reported an association with performance of more than one CAT measure are included in each relevant category; thus, the number of effects is 43.

Table 2.3

Effects on Performance of Cognitive, Physiological, and Dichotomous CAT Variables

CAT Variable	Reference Number		Number of Effects	Percentage of Effects Supporting the Association			Sum Code
				Positive	Negative	None	
Cognitive	-	2, 3, 6, 7, 9, 13, 14, 15, 16, 17, 18, 22, 27, 28, 29, 30	17	76	0	24	++
Physiological	-	1, 8, 14, 17, 20, 23, 24, 26, 27, 28	12	67	0	33	++
Dichotomous	-	4, 5, 10, 11, 12, 13, 15, 19, 20, 21, 25	15	67	7	27	++

Note. Percentages are rounded to integers so do not always total 100. The "Sum Code" was adapted from Sallis, Prochaska, and Taylor (2000): "0" indicates that 0 – 33% of the supported an association, "?" indicates that 34 – 59% of the studies supported the association, and "+" indicates that 60% or more of the studies supported the association. Codes are doubled ("??", "00", or "++" when four or more studies supported the association/lack of association). For the "Reference Number" column coding, please consult the corresponding column in table 2.1.

Sixteen studies reported 17 analyses that examined the association between a cognitive CAT measure and performance. Thirteen analyses (76%) found a statistically significant effect favouring a challenge state, with two effects contingent on interactions (Study 1, White, 2008; Chalabaev et al., 2012). Four analyses found no significant effect (Chalabaev et al., 2009; Laborde et al., 2015; Quigley et al., 2002; Rith-Najarian et al., 2014). Of the six effect sizes reported, three were small (Chalabaev et al., 2012; Moore, Young, Freeman, & Sarkar, 2018; Study 1, Moore, Wilson et al., 2013), two were medium (Study 1, O'Connor et al., 2010; Schneider, 2004), and one was large (Vine et al., 2015). The majority of the cognitive CAT indices used self-report items from Tomaka and colleagues' (1993) cognitive appraisal ratio or Schneider's (2008) stressor appraisal scale to create demand and resource evaluation scores. These scores were combined into a ratio (i.e., demands divided by resources; e.g., Quigley et al., 2002) or a difference score (i.e., resources minus demands; e.g., Chalabaev et al., 2012). However, some studies used single-item measures that assessed the degree to which participants felt challenged or threatened (e.g., Turner et al., 2012).

Eleven studies reported 12 analyses that examined the association between a physiological CAT measure and performance. Eight (67%) found that a challenge cardiovascular response was associated with better performance than the threat response (Blascovich et al., 2004; Moore et al., 2018; Scheepers, 2017; Scholl et al., 2017; Seery et al., 2010; Turner et al., 2013; Studies 1 and 2, Turner et al., 2012). Four analyses found no significant effect (Mendes et al., 2007; Rith-Najarian et al., 2014; Seery et al., 2010; Vine et al., 2013). Of the 10 effect sizes reported, five were small (Blascovich et al., 2004; Moore et al., 2018; Scheepers, 2017; Scholl et al., 2017; Seery et al., 2010), and five were medium (Scholl et al., 2017; Studies 1 and 2, Turner et al., 2012). The physiological CAT index comprised a sum score of the changes in CO and TPR from baseline to a post-instruction (or manipulation) period. These changes were determined by using difference scores in all studies in the "Physiological" group. However, two studies in the "Dichotomous" group used residualised change scores (i.e., standardised residuals of a regression of post-instruction on baseline values, to control for differences in baseline values) to create the index (e.g., Moore et al., 2015; Moore et al., 2014). Both approaches typically weighted TPR reactivity negatively, so that a greater value on the summed CAT index was more reflective of a challenge state. Finally, the timing and duration of physiological data differed between studies. For example, some studies recorded five minutes of baseline data and one minute after giving task instructions, although they often only used the final minute of the baseline period in the analyses (e.g., Moore et al., 2014). Other studies measured five minutes of baseline data and two minutes of reactivity data during the task, using mean values of the entire time periods (e.g., Blascovich et al., 2004).

Only 11 studies included both physiological and cognitive CAT indices, and only three of these studies reported associations with performance for both indices[3] (Moore et al., 2018; Rith-Najarian et al., 2014; Vine et al., 2013). Moore and colleagues (2018) found that both the cognitive and physiological CAT measures were related to performance. Rith-Najarian and colleagues (2014) found that neither measure was related to performance. Vine and colleagues (2013) found that only the cognitive CAT measure was related to performance, with a challenge state linked with better performance. Further, only three of the studies that computed both cognitive and physiological CAT measures provided a correlation between the two indices[4] (Moore et al., 2018; Turner et al., 2013; Vine et al., 2013). Moore et al. (2018; r = .19) and Turner et al. (2013; r = .21) found no significant correlation, whereas Vine et al. (2013) found a

[3] Chalabaev et al.'s (2009) study is not listed here despite reporting performance analyses for the cognitive and physiological variables (i.e., CO and TPR reactivity). This is because the physiological CAT variables were not combined into a single CAT index, which violated the inclusion criteria. However, it is noteworthy that this analysis did find challenge reactivity to be associated with better performance, supporting the contentions of the BPSM.
[4] Two other studies provided associations between cognitive and physiological variables, but did not use a single physiological CAT index (Turner et al., 2012; Quigley et al., 2002). Turner et al. (2012) did not find any significant correlations, although the coefficients were consistent with the BPSM in terms of direction. Quigley et al. (2002) found a marginally significant association between cognitive CAT and CO, but not between cognitive CAT and TPR.

significant correlation during the baseline test (*r* = .32), but not the pressurised test (*r* = -.11).

Fifteen studies created dichotomous groups, which were confirmed with a manipulation check using a cognitive and/or physiological CAT measure. Ten (67%) studies found that the challenge group significantly outperformed the threat group (Study 2, Feinberg & Aiello, 2010; Studies 1 and 3, Gildea et al., 2007; Moore et al., 2012; Moore et al., 2014; Moore et al., 2015; Study 2, Moore, Wilson et al., 2013; Study 2, O'Connor et al., 2010; Scheepers, 2017), with one effect contingent on an interaction (O'Connor et al., 2010). Furthermore, Feinberg and Aiello (2010) reported three significant interaction effects between CAT instructions and experimenter presence. However, they did not report whether challenge was related to better performance than threat in any of the two experimenter presence conditions, comparing challenge with challenge, and threat with threat across the two conditions instead. Four studies found no significant effect (Study 4, Feinberg & Aiello, 2010; Study 2, Gildea et al., 2007; Sammy et al., 2017; Turner et al., 2014), and one study found that participants in the threat condition outperformed those in the challenge condition, although it should be noted that the manipulation check in this study was only marginally significant (Study 1, Feinberg & Aiello, 2010). Of the 16 effect sizes reported, six were small (Study 2, Gildea et al., 2007; Moore et al., 2014; Moore et al., 2018; Study 2, O'Connor et al., 2010; Scheepers, 2017), seven were medium (Study 3, Gildea et al., 2007; Moore et al., 2012; Study 2, Moore, Wilson et al., 2013; Schneider, 2004; Turner et al., 2014), and three were large (Study 1, Feinberg & Aiello, 2010; Study 1, Gildea et al., 2007; Moore et al., 2015).

2.4.3.2 Effects of CAT states on cognitive and behavioural task performance. The performance tasks varied across studies, but could be placed into two main categories: Cognitive and behavioural. Table 2.4 lists the studies in each category and their corresponding results.

Table 2.4

Effects of CAT States on Cognitive and Behavioural Task Performance

Performance Outcome	Reference Number		Number of Effects	Percentage of Effects Supporting the Association: Positive	Negative	None	Sum Code
Cognitive	-	2, 3, 4, 5, 6, 7, 8, 16, 20, 21, 22, 23, 24, 26, 30	23	65	4	30	++
Behavioural	-	1, 9, 10, 11, 12, 13, 14, 15, 17, 18, 19, 25, 26, 27, 28, 29	19	84	0	16	++

Note. Percentages are rounded to integers so do not always total 100. The "Sum Code" was adapted from Sallis et al. (2000): "0" indicates that 0 – 33% of the supported an association, "?" indicates that 34 – 59% of the studies supported the association, and "+" indicates that 60% or more of the studies supported the association. Codes are doubled ("??", "00", or "++" when four or more studies supported the association/lack of association). For the "Reference Number" column coding, please consult the corresponding column in table 2.1.

Twenty studies reported 23 effects involving cognitive performance outcomes, of which eight were mathematical (e.g., serial subtraction task; Kelsey et al., 2000). Examples of other tasks included Stroop (Study 1, Turner et al., 2012), and word-finding (Mendes et al., 2007) tasks. Fifteen (65%) analyses found that a challenge state was associated with superior performance, although two of these effects were contingent on an interaction with another variable (Chalabaev et al., 2012; Study 1, White, 2008). Seven effects were not significant, and one analysis found that participants performed significantly better in the threat condition (Study 1, Feinberg & Aiello, 2010). Of the 15 effect sizes, four were small (Chalabaev et al., 2012; Scholl et al., 2017; Seery et al., 2010), nine were medium (Study 3, Gildea et al., 2007; Schneider, 2004; Scholl et al., 2017; Studies 1 and 2, Turner et al., 2012), and two were large (Study 1, Feinberg & Aiello, 2010; Study 1, Gildea et al., 2007).

Nineteen effects involved behavioural tasks such as golf putting (Moore et al., 2012; Moore et al., 2015; Study 2, Moore, Wilson et al., 2013), cricket batting (Turner et al., 2013), flight simulation (Vine et al., 2015), and a medical selection practical (Roberts et al., 2015). Sixteen (84%) effects favoured a challenge state, with one effect qualified by an interaction with another variable (Study 2, O'Connor et al., 2010). Three effects were not significant (Rith-Najarian et al., 2014; Sammy et al., 2017; Turner et al., 2014). Of the 15 effect sizes reported, six were small (Blascovich et al., 2004; Moore et al., 2014; Study 1, Moore, Wilson et al., 2013; Moore et al., 2018; Study 2, O'Connor et al., 2010), seven were medium (Moore et al., 2012; Study 2, Moore, Wilson et al., 2013; Study 1, O'Connor et al., 2010; Turner et al., 2014; Studies 1 and 2, Turner et al., 2012), and two were large (Moore et al., 2015; Vine et al., 2015).

2.4.3.3 Effects of CAT states on performance within different research designs. Four types of research designs were used: (1) experiments that directly manipulated CAT states (explicitly targeting CAT states), (2) experiments that indirectly manipulated CAT states (targeting another variable, including putative CAT antecedents), (3) correlational studies, and (4) quasi-experiments. Table 2.5 lists the studies grouped by research design. Although the "dichotomous" group in Table 2.3

shares some studies with the "experimental (direct)" and "quasi-experimental" groups, the research questions pertaining to Table 2.3 and Table 2.5 are different. Table 2.3 is about the type of CAT measure and analysis, whereas Table 2.5 is about the type of research design.

Table 2.5

Effects of CAT States on Performance within Different Research Designs

Research Design	Reference Number		Number of Effects	Percentage of Effects Supporting the Association: Positive	Negative	None	Sum Code
Experimental (direct)	-	4, 10, 13, 15	6	67	17	17	++
Experimental (indirect)	-	2, 3, 8, 9, 11, 12, 15, 19, 20, 25, 30	12	67	0	33	++
Correlational	-	1, 6, 7, 13, 14, 16, 17, 18, 22, 23, 24, 26, 27, 28, 29	18	78	0	22	++
Quasi-Experimental	-	5, 21	4	100	0	0	++

Note. Percentages are rounded to integers so do not always total 100. The "Sum Code" was adapted from Sallis et al. (2000): "0" indicates that 0 – 33% of the supported an association, "?" indicates that 34 – 59% of the studies supported the association, and "+" indicates that 60% or more of the studies supported the association. Codes are doubled ("??", "00", or "++" when four or more studies supported the association/lack of association). For the "Reference Number" column coding, please consult the corresponding column in table 2.1.

Six studies reported experiments that directly manipulated participants into CAT states by framing the task instructions consistent with either a challenge or threat state (i.e., perceptions of task demands and personal coping resources). Four (67%) studies found that participants in the challenge group performed significantly better than those in the threat group (Study 2, Feinberg & Aiello, 2010; Moore et al., 2012; Study 2, Moore, Wilson et al., 2013), although one effect was qualified by an interaction (Study 2, O'Connor et al., 2010). One study found no significant effect (Study 4, Feinberg & Aiello, 2010), and one study found that the threat group outperformed the challenge group (Study 1, Feinberg & Aiello, 2010). Of the five effect sizes, one was small (Study 2, O'Connor et al., 2010), three were medium (Moore et al., 2012; Study 2, Moore, Wilson et al., 2013), and one was large (Study 1, Feinberg & Aiello, 2010).

Twelve studies reported experiments that indirectly manipulated CAT states by manipulating another variable such as resource appraisals (Turner et al., 2014), perceived effort and support (Moore et al., 2014), or interpretations of physiological arousal (Moore et al., 2015), and obtained different CAT responses between groups. Eight (67%) studies found that a challenge state was associated with superior performance, although one effect was contingent on an interaction (O'Connor et al., 2010). Four studies found no significant effect (Chalabaev et al., 2009; Mendes et al., 2007; Sammy et al., 2017; Turner et al., 2014). Of the six effect sizes reported, three were small (Chalabaev et al., 2012; Moore et al., 2014; Scheepers, 2017), two were medium (Study 1, O'Connor et al., 2010; Turner et al., 2014), and one was large (Moore et al., 2015).

Sixteen studies used a correlational design, correlating either a cognitive or physiological CAT measure with performance. Of the 18 effects in this group, 14 (78%) showed a significant association between CAT and performance, with a challenge state related to better performance. Four analyses found no significant association (Laborde et al., 2015; Quigley et al., 2002; Rith-Najarian et al., 2014; Seery et al., 2010). Of the 12 effect sizes reported, five were small (Blascovich et al., 2004; Moore et al., 2018; Scholl et al., 2017; Seery et al., 2010), six were medium (Study 2, Moore, Wilson et al., 2013; Scholl et al., 2017; Studies 1 and 2, Turner et al., 2012), and one was large (Vine et al., 2015).

Finally, four studies used a quasi-experimental approach by dividing the sample into CAT groups based on scores on a cognitive CAT measure. All four (100%) studies found that participants in the challenge group performed significantly better than those in the threat group (Gildea et al., 2007; Schneider, 2004). Of the six effect sizes reported, one was small (Study 2, Gildea et al., 2007), four were medium (Study 3, Gildea et al., 2007; Schneider, 2004), and one was large (Study 1, Gildea et al., 2007).

2.5 Discussion

For over two decades, the BPSM of CAT states has been used as a framework to understand variations in cognitive, physiological, and behavioural responses in motivated performance situations (Blascovich & Mendes, 2000). The aim of this systematic review was to examine the relationship between CAT states and performance, and the consistency of this relationship across different CAT measures, performance tasks, and research designs. In 28 (74%) of the 38 studies, a challenge state was related to better performance. Based on statistical significance, the relationship between CAT states and performance was relatively consistent across different measures of CAT states (cognitive vs. physiological vs. dichotomous), performance outcomes (cognitive vs. behavioural), and research designs (direct experimental vs. indirect experimental vs. correlational vs. quasi-experimental), although there were few studies in the direct experimental group.

The common finding that individuals who exhibited a challenge state outperformed individuals who displayed a threat state, supports the predictions of the BPSM and holds relevance for sports psychologists, coaches, business managers, educators, and other professionals interested in optimising human performance.

The beneficial effect of a challenge state was generally consistent across different CAT measures (i.e., cognitive vs. physiological vs. dichotomous). As such, the findings support the prediction of the BPSM that CAT states occur on both a cognitive (i.e., underlying demand/resource evaluations) and physiological (i.e., accompanying cardiovascular responses) level, and influence performance. However, it is noteworthy that studies including the relationships between both CAT measures and performance found an inconsistent pattern (e.g., Moore et al., 2018; Rith-Najarian et al., 2014; Turner et al., 2013), implying that more research is needed to compare the two measures as predictors of performance. In addition, although the BPSM predicts that different demand and resource evaluations lead to distinct physiological responses (Blascovich, 2008a), only three studies included both cognitive and physiological CAT measures and reported correlations among these variables (Moore et al., 2018; Turner et al., 2013; Vine et al., 2013). Weak to moderate correlations were reported in these studies, raising questions about whether demand and resource evaluations trigger distinct cardiovascular responses, as proposed by the BPSM (Blascovich, 2008a). Indeed, the wider BPSM literature has also demonstrated weak to moderate links between cognitive and physiological markers of CAT (e.g., Zanstra, Johnston, & Rasbash, 2010).

Studies that used a single cognitive measure of CAT states to dichotomise individuals into CAT groups (e.g., via a median split) also tended to support the superiority of a challenge state (e.g., Gildea et al., 2007). However, dichotomising CAT states is incongruent with the notion that they represent opposite ends of a single bipolar continuum (Blascovich & Mendes, 2000). Further, dichotomising a sample with a median split could lead to problems like loss of statistical power and difficulty in comparing results between studies due to the different cut-off points employed (Altman & Royston, 2006). Researchers should therefore consider whether it is appropriate to dichotomise CAT measures and, if so, ensure that the study has sufficient power.

This review revealed notable diversity in the recording and calculation of cognitive and physiological CAT measures. For instance, both single and multiple self-report items assessed demand and resource evaluations (Schneider, 2008; Tomaka et al., 1993; Turner et al., 2013). In addition, responses to these items were used to calculate a ratio (i.e., demands divided by resources; e.g., Moore et al., 2012), or difference (i.e., resources minus demands; e.g., Moore, Wilson, et al., 2013) score. Moreover, CO and TPR were reported as reactivity (e.g., Blascovich et al., 2004) or residualised change scores (e.g., Moore et al., 2012). These values were often calculated by averaging across different durations and time periods (e.g., final minute of baseline and first minute after receipt of task instructions, Moore et al., 2014; or final two minutes of baseline and first two minutes of the task itself, Blascovich et al., 2004). The justifications for these variations were not always clearly articulated and should be made more explicit in future research.

Although these variations did not appear to impact the findings, future research would benefit from adopting a more consistent approach in CAT measurement to facilitate the synthesis of evidence across studies. If studies adopt different methods to measure CAT states, it is unclear whether the observed relationships are due to CAT states themselves or the idiosyncratic measurement processes (e.g., because self-report was employed rather than cardiovascular indices or a ratio vs. a difference score). Although we encourage future research to contrast the different ways of measuring CAT states to empirically identify the optimal approach, we make the following recommendations based on the justifications provided in the current literature. Researchers should use both cognitive evaluations and cardiovascular responses to measure CAT states, and further examine their relationship and respective effects on performance. Given the limitations associated with single-item scales (e.g., lower relative precision than multi-item scales; McHorney, Ware, Rogers, Raczek, & Lu, 1992), multi-item measures of demand and resource evaluations should be employed (e.g., Schneider, 2008). The scores from these items should then be used to calculate a difference score, as ratio scores have been discouraged due to their highly nonlinear distribution (Vine et al., 2013). When measuring the physiological indices of CAT states (i.e., CO and TPR reactivity), researchers should use comparable time periods and indices. To ensure true resting values are obtained, researchers should use the final minute of the baseline period (Sherwood, Allen, Kelsey, Lovallo, & van Doornen, 1990). Further, given the dynamic nature of CAT states (i.e., reappraisal; Blascovich, 2008a), researchers should utilise the first minute after task instructions or of task exposure. While most research has employed difference scores rather than residualised change scores, we recommend that researchers consult guidelines and use the approach most suitable for their data (e.g., Burt & Obradovic, 2013). Finally, CO and TPR reactivity should be combined into a single CAT index, which is more in keeping with the unidimensional nature of CAT states, increases reliability, and simplifies analyses (Seery et al., 2010).

The risk of bias assessment showed that random sequence generation, incomplete outcome data, other sources of bias, blinding of outcome assessment, incomplete outcome data, confounding variables, and intervention (exposure) measurement exhibited a low risk of bias across most studies. Allocation concealment, blinding of participants and personnel, blinding of outcome assessment, selection of participants, and selective reporting often exhibited an unclear risk of bias. As only three studies were rated as high risk of bias, the body of evidence appears to be of adequate quality overall, but the findings highlight the importance of considering and reporting potential risks in future studies. For example, researchers should minimise missing physiological and outcome data, ensure that performance assessors are naive to CAT data, and provide information about allocation concealment, blinding of participants, personnel and outcome assessment, and selective reporting.

Based on statistical significance, there was a relatively consistent relationship between CAT states and performance on behavioural and cognitive tasks. The notable difference in support for cognitive vs. behavioural tasks (see Table 2.3) could have been influenced by the included and excluded studies. First, although Chalabaev et al. (2009) found that greater CO reactivity and lower TPR reactivity were associated with better cognitive performance separately, the review excluded this study as no single physiological CAT index was reported. Second, Feinberg and Aiello’s (2010) three studies that manipulated participants into CAT groups using verbal instructions, found inconsistent effects for CAT states on performance, one of which involved an only marginally significant manipulation check. As well as being inconsistent with the notion that CAT states are a continuum (Blascovich & Mendes, 2000), this approach averages data across CAT groups and individuals who were not successfully manipulated into the required state might have attenuated the results (i.e., individuals in the challenge group displaying a threat state, and vice versa; Turner et al., 2013). As such, the weaker effect on cognitive outcomes might have been caused by other confounding statistical and methodological issues.

Studies that directly manipulated CAT states provided support for the superiority of a challenge state, although only six studies utilised such a design. Four studies found that the challenge group outperformed the threat group (Study 2, Feinberg & Aiello, 2010; Moore et al., 2012; Moore, Wilson et al., 2013; O'Connor et al., 2010), and two studies reported null or contradictory results (Studies 1 and 4, Feinberg & Aiello, 2010). Issues such as the strength and effectiveness of the CAT manipulation instructions (as well as the limitations noted above) might explain the heterogeneous results among Feinberg and Aiello's (2010) studies. For example, Feinberg and Aiello read instructions aloud to participants, whereas Moore et al. (2012, 2013) delivered standardised instructions from memory more directly to participants. Researchers employing experimental designs should report the methods used to manipulate participants into CAT states and use both cognitive and physiological CAT measures as manipulation checks, as the two measures could yield divergent results.

Although two theoretical models (Jones et al., 2009; Vine et al., 2016) have proposed several potential mechanisms through which CAT states might influence performance, only three studies included in the review explicitly tested mediation (Moore et al., 2012; Moore, Wilson et al., 2013 study 2; Vine et al., 2013). Of these studies, only one study reported statistically significant mediation (Moore et al., 2012), with the findings suggesting that CAT states influenced golf-putting performance primarily via kinematic variables and not through emotional, attentional, or physiological pathways.

Despite this limited evidence for significant mediating processes, studies have reported that CAT states are associated with different emotional, attentional, and physiological responses, with a challenge state linked with less cognitive anxiety, more optimal visual attention, and less muscle activity (Moore et al., 2012; Moore, Wilson et al., 2013 study 2; Vine et al., 2013). It is vital for research to continue exploring these and other potential underlying mechanisms to better understand how a challenge state facilitates performance. In particular, research should test the attentional mechanisms outlined by Vine et al. (2016), and examine whether a threat state increases the influence of the stimulus-driven system and draws attention away from task-relevant to less relevant (and potentially negative) stimuli, resulting in suboptimal performance.

Several issues emerged as limitations to the present review. First, a meta-analysis may have provided additional information about the strength of the relationship between CAT states and performance. However, this was not feasible due to the substantial variability in methodologies adopted across studies. The variability across studies also hindered the ability to clearly delineate how strongly the effects were influenced by the CAT measure, task, or research design. Second, as this review only included published studies, publication bias might have influenced its results. Third, the sum codes used in Tables 2.3, 2.4, and 2.5 (adopted from Sallis et al., 2000) use arbitrary cut-off points and refer to patterns of statistical significance, which do not take into account effect sizes. Finally, while the research team categorised tasks as either cognitive or behavioural, many tasks required both cognitive input and behavioural execution. For example, golf putting requires cognition to determine the optimal direction and behavioural control to execute the motor skill.

This review highlights key directions for future research. Given that a challenge state facilitates performance, it is important to identify factors that elicit a challenge state to aid the development of theory and effective interventions. While some antecedents proposed by the BPSM (e.g., required effort and support; Moore et al., 2014) and TCTSA (e.g., control, self-efficacy, and achievement goals, Turner et al., 2014) have been investigated, research should examine other possible antecedents (e.g., danger, uncertainty, familiarity, knowledge, skills, abilities; Blascovich, 2008a). Further, although some interventions have received attention (e.g., arousal reappraisal, Moore et al., 2015), research should examine other interventions aimed at promoting a challenge state. Finally, the longitudinal (and likely reciprocal) relationship between CAT states and performance should be explored.

2.5.1 Conclusion

To conclude, a challenge state was related to better performance than a threat state in 74% of studies. The quality of the included studies was generally good, although the risk of bias assessment identified some areas for improvement (e.g., minimise data loss). This association between CAT states and performance was relatively consistent across cognitive, physiological, and dichotomous CAT variables; cognitive and behavioural tasks; and direct experimental, indirect experimental, correlational, and quasi-experimental designs. Future research would benefit from a more consistent approach to CAT measurement (e.g., multi-item self-report measures of cognitive evaluations), to reduce ambiguity and aid the synthesis of results across studies. Furthermore, researchers should develop challenge-promoting interventions to optimise the performance of individuals across a range of domains (e.g., sport, academia, business, and medicine).

Table 2.6

Summary of Chapter 2 and Preview of Next Chapter

Chapter	Aim	Findings
2	To systematically review the relationship between CAT states and performance in the published literature.	A challenge state was related to better performance than a threat state in 74% of studies. The association was consistent across CAT variables, outcome tasks, and research designs.
	Rationale for next chapter Given the evidence for the superiority of a challenge state, research should elucidate whether (and to what extent) CAT states vary as a function of differences between persons, situations, or interactions thereof. This research could then pave the way for potential challenge-promoting interventions.	
Chapter	Aim	Findings
3	To partition the variance in CAT states into personal (person), situational (task, week), and interaction components.	

Chapter 3

Examining the Variance of Challenge and Threat States across Tasks and Time Points

3.1 Abstract

Although a challenge state has been associated with better performance than a threat state, no previous research has explored the generalisability of challenge and threat states across people and situations. Also, the cardiac reactivity indexing task engagement has not been well-studied across repeated measures. Thus, this study aimed to explore variance components of challenge and threat states and cardiac reactivity indexing task engagement across repeated measures. Cognitive and cardiovascular indicators of challenge and threat states were measured in 30 participants performing four tasks on three time points (separated by one week each). Variance components analyses decomposed total variances into person, task, week, and their two-way interaction components. Significant person components were found on cognitive and cardiovascular challenge and threat variables (explaining 16-40% of the variance), whereas significant person by task and person by week interaction components were only consistently found on cognitive variables (explaining 6-13% of the variance). Results also indicated that task engagement-related cardiac reactivity was relatively more stable over time than postural stressor-related cardiac reactivity. In sum, the present study presented novel insights into the variance of challenge and threat states, which may guide applied research toward person- or person by situation-based interventions. The results also indicate that task engagement may be relatively stable over time across repeated measures in motivated performance situations.

3.2 Introduction

Cognitive and cardiovascular indicators of CAT states have been well-researched regarding their association with performance, where a challenge state was generally found to be superior to a threat state (see Behnke & Kaczmarek, 2018; chapter 2). However, CAT states have not been well-researched regarding their dynamic nature across repeated measurements, which was highlighted as a gap in the literature in chapters 1 and 2. Thus, there is less consensus in the literature about the dynamic nature of CAT states than about the relationship between CAT states and performance. Some researchers have proposed that there are trait-like tendencies to experience a challenge or a threat state in motivated performance situations (e.g., Tomaka et al., 2018). Others have found considerable within-subject variation in CAT states (e.g., Trotman et al., 2018). However, the lack of consensus in the literature is mostly not due to conflicting evidence, but due to a lack of evidence regarding both personal and situational factors in CAT states. To my knowledge, no previous study has examined how CAT states vary across persons, tasks, and time points. Therefore, the main aim of the present study was to determine the extent to which CAT states vary as a function of personal, situational, and person-by-situation interactional factors across different tasks and repeated measurements.

Although the BPSM describes CAT *states*, it does not exclude the possibility of an overarching CAT trait variable that may explain individual tendencies that are stable over time. Indeed, recently a questionnaire was developed to assess stable individual differences in CAT evaluations consistent with a BPSM perspective (Tomaka et al., 2018). Furthermore, Blascovich and colleagues (2004) found that baseball and softball players who exhibited a challenge state during a pre-season speech about their sport performed better throughout the subsequent season than those who exhibited a threat state. Given the diffuse performance outcome (average season performance) and the large time span of this study (CAT states being measured pre-season, but performance depending in part on late-season scores), it seems likely that the CAT states measured in the pre-season speech reflected stable dispositions and were similar to those experienced in competitions throughout the season. More support for stable CAT tendencies comes from Dienstbier's (1989) work on physiological toughness, which suggested that personality factors correlate with physiological toughness patterns. These physiological toughness patterns (i.e., differential cardiovascular and hormonal responses) were a key influence on the physiological predictions of the BPSM (Blascovich, 2008a; Tomaka et al., 1993). Hence, some evidence exists to suggest a trait component to CAT states, although the topic has not been widely studied.

On the other hand, it seems intuitive that specific situations (e.g., an unexpected extremely strong or weak performance of an opponent) would have the potential to elicit a challenge or a threat state in most individuals. Indeed, some research supports the idea of situational determinants of CAT states, as it found a public speaking task to be more threatening than a pressurised competition in a car racing video game (Trotman et al., 2018). In particular, the public speaking task elicited greater demand and lesser coping resource evaluations, more anxiety and debilitative anxiety interpretations, and lower perceived control. Furthermore, Mendes and colleagues (2002) found that when the situation varied as to presenting participants with different partners on a cooperative tasks, CAT states varied as well. For example, when paired with Black (versus White) or socioeconomically disadvantaged (versus advantaged) confederates, participants exhibited cardiovascular responses more indicative of a threat state These findings suggest that situational determinants may indeed play a role in CAT states, although again, this research topic has not received much explicit attention yet.

Studies have also hinted at the existence of person by situational interactions, as a personal variable interacted with a situational variable to predict CAT states (Blascovich et al., 1999). In particular, participants who performed a well-learned task in front of an audience exhibited a relative challenge cardiovascular pattern, whereas those who performed a novel task exhibited a relative threat pattern. However, these differences were not replicated when participants performed alone, indicating a person by situation interaction in CAT states. Thus, person by situation interaction effects might explain variance in CAT states, but previous research has not yet elucidated the exact proportion in relation to personal and situational factors.

To decompose the variance in cognitive and cardiovascular indicators of CAT states, the present study used a generalisability theory framework (Brennan, 2011; Shavelson & Webb, 2006). Generalisability theory is a suitable and widely used approach to determine how large an influence can be attributed to personal (i.e., trait), situational (i.e., state), and interaction (i.e., trait x state) components (Lakey, 2016). For example, it has been applied to the context of social support (Lakey, Lutz, & Scoboria, 2004; Rees, Freeman, Bell, & Bunney, 2012), interpersonal perceptions (Kenny, West, Malloy, & Albright, 2006), memory performance (Gross et al., 2015), and appraisals of work stressors in police officers (Lucas et al., 2012). In the latter, personal (14-15%), situational (18-19%), and person by situation interactional (38-41%) components were found, indicating that there are individual differences in how officers generally appraise work stressors, certain differences between stressors that are stable across officers, and differences between stressors that are different between officers. As the analysis of CAT states provides a similar context to Lucas and colleagues' (2012) analysis of work stressor appraisals, a generalisability approach might also provide fundamental insights into the variability of CAT states as a function of person, task, time point, and interaction components. The present study first applied this method to the measurement of CAT states.

Using a generalisability theoretical approach to partition the variance in CAT states across tasks and measurements could provide insights carrying practical implications for sport professionals interested in optimising performance, as a challenge state has been shown to be superior to a threat state in terms of performance (see Behnke & Kaczmarek, 2018; chapter 2). In particular, identifying the main sources of variation in CAT states could guide and facilitate the development of effective challenge-promoting interventions. For example, if CAT states were found to vary mainly as a function of personal factors, challenge-promoting interventions should target these personal factors to help those individuals who habitually experience a threat state. Conversely, if CAT states were found to vary mainly as a function of situational factors, then sport professionals would want to target those situational factors that provoke a threat state in their athletes. Finally, it could be that CAT states vary as a function of interactions between the person and the situation. That is, some athletes would experience more challenge than others on one task, but this pattern might be reversed on another task. Partitioning the variance into personal, situational, and interaction components bears relevance for sport professionals because even the best person-focused intervention would be conducted in vain if CAT states were to vary largely as a function of situational factors (and vice versa). Thus, a generalisability analysis of CAT states could guide the development of effective challenge-promoting interventions.

As the BPSM specifies task engagement as a prerequisite for CAT states to be analysed, a research question of practical relevance would be whether the cardiovascular response used as a proxy for task engagement (i.e., HR reactivity) is stable or varies over time. If a blunting of this cardiovascular response were to occur, it would provoke the question of whether this blunting is due to a decrease in task engagement, or a cardiovascular habituation effect over time. One study has compared the cardiovascular response indicative of task engagement across four mental arithmetic tasks performed in one session (Kelsey, Soderlund, & Arthur, 2004). It did indeed find an attenuation of HR reactivity across tasks. Interestingly, the cardiovascular adaptation was partly reversed by evaluative observation, which some participants were exposed to in the third task. Another study examined HR reactivity between a public speaking task and a video game competition. Both tasks were performed in the same testing session and there were no differences in HR reactivity (Trotman et al., 2018). However, to my knowledge no previous study has recorded and compared HR reactivity in participants performing different tasks, and repeating the same tasks on different days. Comparing such data would provide important insights into the stability of the task engagement cardiovascular response. Particularly, it would allow to test whether Kelsey and colleagues' (2004) findings generalise to the same task being performed on different days, or whether results across different time points would resemble those of Trotman and colleagues' (2018).

In addition to comparing task engagement-related HR reactivity between various tasks and time points, this study also included a measure of postural challenge-related HR reactivity (the HR response to quickly standing up and sitting down again) to compare with the task engagement-related measure over time. The rationale for this was to provide more conclusive evidence regarding the change of task engagement over time.

For example, if task engagement-related HR reactivity were to decrease, but postural challenge-related HR reactivity were to remain stable over time, then the decrease could likely be attributed to an actual decrease in task engagement. However, if both task engagement-related and postural challenge-related HR reactivity were to decrease equally strongly over time, then the decrease might reflect a general cardiovascular habituation effect, rather than decreased task engagement.

The current study examined whether differences in cognitive CAT evaluations and cardiovascular CAT responses can be attributed to differences between persons, situations (i.e., tasks or time points), or interactions between these factors. In particular, we hypothesised at least one situational (state) component (task and/or time point), a person (trait) component, and at least one interaction component (H1). The secondary aim was to examine whether the cardiovascular response used as a proxy for task engagement varies by time points. While this is not theoretically relevant to the partitioning of CAT states, it is practically relevant to CAT researchers interested in collecting multiple measures of CAT states (e.g., on multiple tasks or in multiple weeks). Therefore, this study also explored whether cardiac reactivity to a psychological stressor (i.e., task engagement) was as stable over time (i.e., task order and week) as cardiac reactivity to a postural stressor (H2). The tertiary aim was to examine the relationship between CAT states and performance, wherein cognitive evaluations and cardiovascular responses consistent with a challenge state were hypothesised to relate to better performance than those consistent with a threat state (H3).

3.3 Method

3.3.1 Participants

The sample consisted of 33 students and staff members at the University of Essex. Two participants dropped out before the second, and one dropped out before the third session, leaving a final sample of 30 participants (28 male, 2 female). Participants' age ranged from 18 to 35 years, with a mean of 23.4 years (*SD* = 4.9). All participants reported being right-handed or ambidextrous.

3.3.2 Materials

3.3.2.1 Cardiovascular data. The Portapres Model-2 (Finapres Medical Systems BV, Amsterdam, the Netherlands) was used to record cardiovascular variables: HR, TPR, and CO. It bases its measurements on the arterial volume-clamp method of Peñáz (1973) and the physiological calibration criteria for the proper unloading of the finger arteries of Wesseling (1996). It also uses a height correction unit to compensate for hydrostatic pressure changes due to movement of the hand. Previous research has used the Portapres in a CAT setting (Moore et al., 2018; Zanstra et al., 2010) and has validated it against the Finapres and the Oxford method, finding it to be accurate, reliable, and cause no more missing data due to artefacts than the Oxford method (Hirschl, Woisetschläger, Waldenhofer, Herkner, & Bur, 1999; Imholz et al., 1993). Data were converted and downloaded with Beatscope version 1.1.

3.3.2.2 Cognitive evaluations. Demand and resource evaluations were assessed with two items commonly used in previous research (e.g., Vine et al., 2013): "How demanding do you expect the upcoming task to be?" for demands, and "How able are you to cope with the demands of the upcoming task?" for resources. Both items were scored on a seven-point Likert scale anchored by *not at all* (1) and *extremely* (7). A cognitive CAT variable was then created by subtracting demands from resources, meaning that possible scores ranged from -6 to 6 and denoted more challenge as values increased.

3.3.2.3 N-back task. The N-Back task (Kirchner, 1958) was administered via a Qualtrics survey, which presented a string of 23 letters (see Appendix A) for five seconds each. Starting at the fourth letter, participants were prompted to indicate (by ticking one of two boxes saying *yes* or *no*) whether the letter shown on the current page was equal to the letter shown three screens earlier (3-back condition). Thus, there were 20 items in total, 10 of them requiring *yes* and 10 of them requiring *no* as the correct answer. Answer choice and time taken to respond was recorded for each item (up to a maximum of five seconds if there was no response).

3.3.2.4 Subtraction task. A Qualtrics survey presented 20 multiple-choice subtraction exercises (see Appendix B) involving the subtraction of a three-digit number from another three-digit number. Exercises were presented separately and sequentially. Each screen presented the correct solution and three false answers in randomised order. Answer choice and time taken to respond was recorded for each item (up to a maximum of ten seconds if there was no response).

3.3.2.5 Bean-bag throwing task. The task consisted of 20 throws. A bean-bag weighing 80 g, measuring 6 x 5 x 5 cm was thrown from a distance of 4 m to a 50 x 50 cm quadratic target on the laboratory floor. Participants scored one point each time the bean-bag came to rest on the target. There was no time limit for this task.

3.3.2.6 Dart-throwing task. The task consisted of 20 throws. Participants threw a Winmau Family Dart Game dart from a distance of 2.4 m toward a Winmau Family Dart Game dartboard. The back of the board was used, which (unlike a traditional dart board) is divided into 9 outer rings and a red bulls-eye area only. Participants were instructed that they needed to throw the dart into the central three areas (“8”, “9”, or the bulls-eye) to score a point. There was no time limit for this task.

3.3.3 Procedure

The study received ethical approval from the University of Essex. The experimenters approached participants in person and through the university e-mail system. Upon entering the lab, participants read an information sheet and provided written informed consent. The information sheet highlighted that four £30 rewards were available for the best performers on each task and that one participant would be randomly drawn to win £40 in each week of the study. After giving informed consent, participants were seated in front of a computer, on which a Qualtrics survey was opened to guide them through the study. On the first day of measurement only, participants provided demographic information. The experimenter then put the Portapres on the left arm of the participant, placing the cuff around the middle finger. In two cases, the participant’s index finger was used instead due to unsuccessful measurements. Participants then went on to the first task. The data collection procedure is graphically summarised in Figure 3.1. The order of the four tasks was randomised for each participant on each measurement occasion (see Figure 3.2).

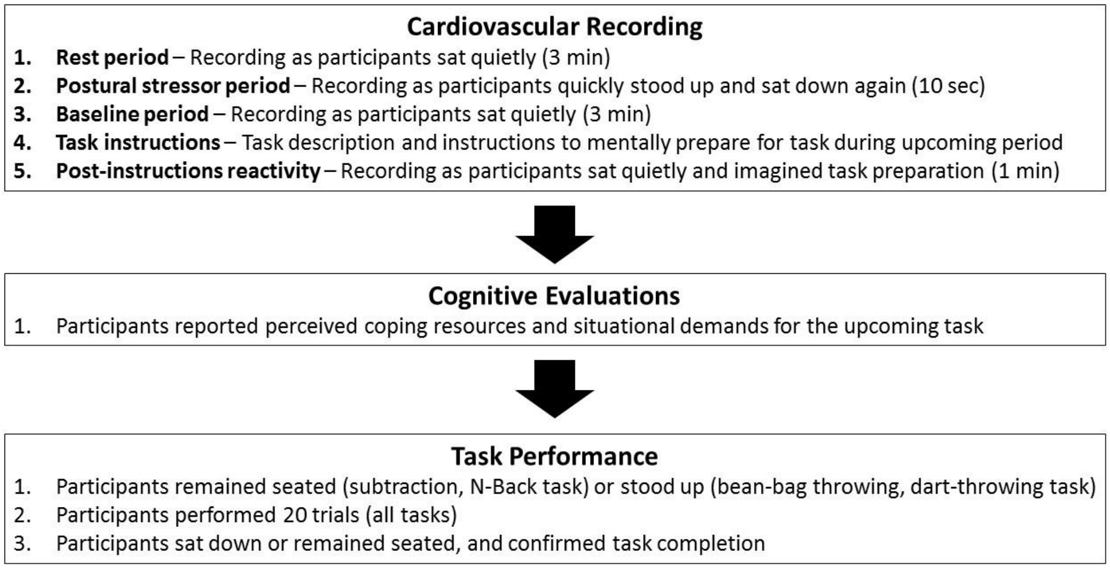

Figure 3.1. Procedure for obtaining cardiovascular, cognitive, and performance data for each task.

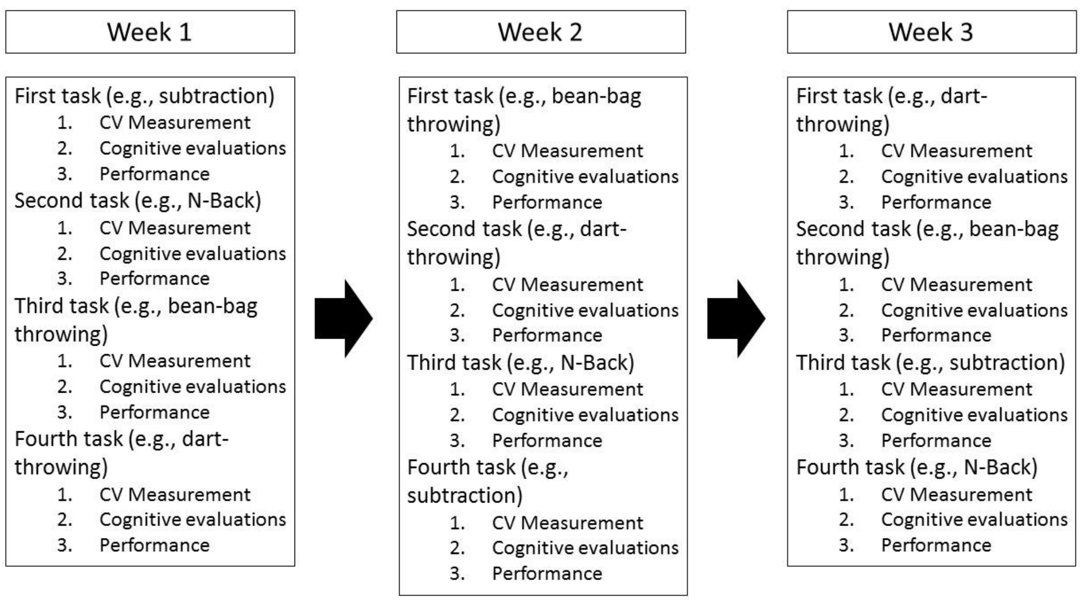

Figure 2.2. Study protocol for data collection sessions.

Before starting each task, cardiovascular responses were recorded throughout four measurement periods: rest period (3 min), postural stressor (10 sec), baseline period (3 min), and post-instructions reactivity period (1 min). The rest period provided the data to be compared against the postural stressor period data. During the postural stressor period, participants were asked to quickly stand up and sit down again. These 10 seconds were used to examine differences in cardiovascular reactivity between participants, as well as differences in reactivity within participants. The baseline period provided the data to be compared against the post-instructions period data. Between the baseline and the post-instructions reactivity period, participants saw a screen displaying instructions for the upcoming task. Other than information about and rules for the upcoming task, this screen reminded participants of the £30 reward for the best performer on the upcoming task, as well as the fact that quicker task completion time would determine a winner between participants with an equal score on the task. By confirming that they had read and understood the task instructions, participants continued to the next screen, which started the post-instructions period. During this period, participants were instructed to sit still for one minute and mentally prepare for the upcoming task. Once the minute had elapsed, participants were asked about their cognitive evaluations. After participants had reported their cognitive evaluations, they started the task. Once the task was completed, the procedure was repeated for the second, third, and fourth task, respectively. After the fourth task, participants were thanked for participating and asked to come back one week later at the same time. The procedure was the same in week 2 and week 3, after which the study was complete.

3.3.4 Statistical Analysis

Mean HR, CO, and TPR values were calculated for the respective last minute of the rest and the baseline period, as well as for the entire postural stressor (10 sec) and post-instructions reactivity period (1 min). Twenty-eight univariate outliers (values more extreme than three standard deviations from the mean; Stevens, 2009) were winsorised to be 1% more extreme than the next non-outlying score[5] (as Shimizu, Seery, Weisbuch, & Lupien, 2011). Baseline CO and TPR values were then regressed on their respective reactivity values with the standardised residuals being saved to create residualised change scores[6]. TPR residualised change scores were then subtracted from CO residualised change scores to create a single cardiovascular CAT index, on which greater values denoted cardiovascular responses more consistent with a challenge state. For additional analyses, the raw differences between reactivity and baseline CO and TPR means were calculated to create raw CO and TPR change scores, respectively.

A Generalised Estimating Equations (GEE) analysis specifying week, order, and measurement period (coding for rest period, postural challenge, baseline, and post-instructions) as within-subjects effects compared mean HR across all weeks, tasks, and measurement periods. Simple contrasts were used to compare groups; the respective last category being the reference category (i.e., week 3, task 4, and post-instructions). To assess task engagement, the main effect for measurement period was examined (i.e., the contrast between baseline and post-instructions). To examine whether HR in the baseline period was significantly different from the rest period (i.e., whether HR returned to resting levels after the postural challenge), the analysis was repeated with rest period as the reference group. All GEE analyses specified an independent correlation structure and used a significance level of $\alpha = .05$.

H1 was tested with six variance components analyses on the outcome variables of DRES, demands, resources, cardiovascular CAT index, raw CO change, and raw TPR

[5] These analyses were conducted separately for each week and each task. In week 1, there were five outliers on task 1, one on task 2, three on task 3, and one on task 4. In week 2, there were three outliers on task 1, three on task 2, two on task 3, three on task 2, two on task 3, and three on task 4. In week 3, there were two outliers on task 1, three outliers on task 2, and two outliers on task 4.
[6] These analyses were conducted separately for each week and each task.

change. Using the restricted maximum likelihood estimation method, variances were partitioned into components for person, task, week, their two-way interactions, and an error term (which is confounded with the highest order interaction). Components were tested for significance by computing a 95% confidence interval and examining whether it excluded zero (Lakey et al., 2004). To do this, the following formula was used, where *x* denotes the respective variance component and *var* denotes the variance of the respective variance component:

$$95\%\ \mathrm{CI}(x \pm \sqrt{var_x} \times 1.96)$$

To assess whether task engagement or postural HR reactivity changed as time progressed (H2), a second GEE analysis was conducted to analyse raw HR change by cardiovascular reference period (postural versus psychological), week, order, and their two-way interaction effects. Differences over time between the two cardiovascular reference periods were explored by examining the reference period by week and reference period by order interaction effects.

To test H3, a third GEE analysis predicted task performance with cognitive CAT, cardiovascular CAT, week, task, and order. Week, task, and order were specified as within-subjects effects.

3.4 Results

Due to equipment problems, cardiovascular data could not be recorded for nine participants on some tasks and/or weeks, which led to 7.5% of total cardiovascular data missing in the analyses. The GEE analysis of baseline and post-instructions HR data found a significant main effect for measurement period (Wald $\chi^2 = 409.89$, $p < .001$). HR was significantly lower in the baseline than in the post-instructions period, indicating sufficient task engagement and thereby permitting the analysis of CAT states [$B = -1.25$, Wald χ^2 = 8.98, p < .01, 95% CI (-2.07, -0.43)]. HR means for the four measurement periods by task order and week are detailed in Table 3.1. Changing the reference category to the rest period revealed that HR was significantly lower in the baseline than in the rest period, suggesting that HR fully recovered after the postural stressor [B = -0.57, Wald χ^2 = 5.04, p = .02, 95% CI (-1.07, -0.07)].

Table 3.1

Estimated Marginal Means for Week by Task Order

	MP	Week 1 *M*	Week 1 *SE*	Week 2 *M*	Week 2 *SE*	Week 3 *M*	Week 3 *SE*	Total *M*	Total *SE*
TO 1									
	RP	77.78	2.16	80.33	2.19	76.79	1.93	78.30	1.88
	PS	89.31	2.02	91.86	1.87	88.32	1.70	89.83	1.62
	BL	77.19	2.18	79.75	2.25	76.20	1.98	77.71	1.92
	PI	78.34	2.15	80.89	2.18	77.35	1.95	78.86	1.88
TO 2									
	RP	77.18	2.07	79.74	2.09	76.20	1.81	77.71	1.76
	PS	88.72	1.93	91.27	1.77	87.73	1.58	89.24	1.50
	BL	76.60	2.09	79.15	2.15	75.61	1.85	77.12	1.81
	PI	77.75	2.07	80.30	2.10	76.76	1.84	78.27	1.78
TO 3									
	RP	76.33	2.05	78.88	2.10	75.34	1.76	76.85	1.74
	PS	87.86	1.93	90.41	1.78	86.87	1.53	88.38	1.49
	BL	75.74	2.08	78.30	2.16	74.75	1.81	76.26	1.80
	PI	76.89	2.06	79.44	2.11	75.90	1.80	77.41	1.77
TO 4									
	RP	75.79	1.99	78.34	2.04	74.80	1.78	76.31	1.70
	PS	87.32	1.84	89.88	1.69	86.33	1.53	87.84	1.41
	BL	75.20	2.00	77.76	2.09	74.22	1.81	75.73	1.74
	PI	76.35	1.99	78.90	2.04	75.36	1.81	76.87	1.72
Total									
	RP	76.77	2.05	79.32	2.08	75.78	1.80	77.29	1.75
	PS	88.30	1.91	90.86	1.75	87.31	1.56	88.82	1.48
	BL	76.18	2.07	78.74	2.14	75.20	1.84	76.71	1.80
	PI	77.33	2.05	79.88	2.09	76.34	1.83	77.85	1.76

Note. Dependent variable: Mean HR (beats per minute). $N = 1335$. MP = Measurement period. TO = Task order. RP = Rest period. PS = Postural stressor. BL = Baseline. PI = Post-instructions.

3.4.1 Variance Components Analyses

The variance components analyses of cognitive evaluations (cognitive CAT, resources, and demands) are detailed in Table 3.2. The analysis of cognitive CAT found significant variance components for the person (explaining 39.6% of the variance), the person by task interaction (explaining 12.1% of the variance), the person by week interaction (explaining 7.1% of the variance), and the error term (explaining 29.3% of the variance). The variance components analysis of resource evaluations found significant variance components for the person (explaining 40.0% of the variance), the person by task interaction (explaining 5.9% of the variance), the person by week interaction (explaining 10.8% of the variance), and the error term (explaining 35.4% of the variance). The variance components analysis of demand evaluations found significant variance components for the person (explaining 39.9% of the variance), the person by task interaction (explaining 13.1% of the variance), the person by week interaction (explaining 8.6% of the variance), and the error term (explaining 28.2% of the variance).

Table 3.2

Variance Components Analyses of Cognitive Evaluations

Source		Component	Percentage of Variance	95% Confidence Interval
Cognitive CAT				
	Person	2.30	39.61	(0.87, 3.73)*
	Week	0.00	0.00	N/A
	Task	0.70	11.98	(-0.49, 1.88)
	Person*Task	0.70	12.05	(0.30, 1.10)*
	Person*Week	0.41	7.11	(0.10, 0.73)*
	Week*Task	0.00	0.00	N/A
	Error	1.70	29.25	(1.34, 2.05)*
	Total	5.84	100	
Resources				
	Person	0.65	40.01	(0.25, 1.05)*
	Week	0.01	0.88	(-0.03, 0.06)
	Task	0.12	7.08	(-0.08, 0.31)
	Person*Task	0.10	5.87	(0.00, 0.19)*
	Person*Week	0.18	10.78	(0.05, 0.30)*
	Week*Task	0.00	0.00	N/A
	Error	0.57	35.37	(0.45, 0.69)*
	Total	1.62	100	
Demands				
	Person	0.96	39.91	(0.36, 1.56)*
	Week	0.00	0.00	N/A
	Task	0.23	9.66	(-0.17, 0.64)
	Person*Task	0.31	13.12	(0.15, 0.48)*
	Person*Week	0.21	8.64	(0.07, 0.35)*
	Week*Task	0.01	0.44	(-0.03, 0.05)
	Error	0.68	28.23	(0.53, 0.82)*
	Total	2.39	100	

Note. Significant variance components are denoted by an asterisk (*). Cognitive CAT: *N* = 355. Resources: *N* = 356. Demands: *N* = 359.

The variance components analyses of cardiovascular variables are detailed in Table 3.3. The analysis of the cardiovascular CAT index found significant variance components for the person (explaining 16.2% of the variance) and the error term (explaining 83.1% of the variance). The analysis of the raw CO change data found significant variance components for the person (explaining 16.0% of the variance), the person by week interaction (explaining 12.1% of the variance), and the error term (explaining 71.6% of the variance). The analysis of the raw TPR change data found a significant variance component for the week by task interaction (explaining 0.4% of the variance) and the error term (explaining 84.7% of the variance).

Table 3.3

Variance Components Analyses of Cardiovascular Variables

Source	Component	Percentage of Variance	95% Confidence Interval
Cardiovascular CAT			
Person	0.42	16.19	(0.09, 0.75)*
Week	0.00	0.00	N/A
Task	0.00	0.00	N/A
Person*Task	0.02	0.67	(-0.25, 0.29)
Person*Week	0.00	0.00	N/A
Week*Task	0.00	0.00	N/A
Error	2.16	83.14	(1.75, 2.57)*
Total	2.60	100	
CO Change (raw)			
Person	0.13	16.04	(0.01, 0.26)*
Week	< 0.01	0.27	(-0.02, 0.02)
Task	0.00	0.00	N/A
Person*Task	0.00	0.00	N/A
Person*Week	0.10	12.08	(0.00, 0.20)*
Week*Task	0.00	0.00	N/A
Error	0.60	71.61	(0.50, 0.71)*
Total	0.84	100	
TPR Change (raw)			
Person	< 0.01	8.46	(0.00, 0.01)
Week	0.00	0.00	N/A
Task	0.00	0.00	N/A
Person*Task	< 0.01	6.49	(0.00, 0.01)
Person*Week	0.00	0.00	N/A
Week*Task	< 0.01	0.40	(0.00, 0.00)*
Error	0.02	84.65	(0.02, 0.03)*
Total	0.03	100	

Note. Significant variance components are denoted by an asterisk (*). $N = 333$.

3.4.2 GEE Analyses of HR Variables

The first GEE analysis of mean HR data found significant main effects for measurement period (Wald $\chi^2 = 384.62$, $p < .001$) and order (Wald $\chi^2 = 16.55$, $p < .001$), as well as a marginally significant effect for week (Wald $\chi^2 = 5.70$, $p = .06$). Parameter estimates for this analysis are summarised in table 3.4. The parameter estimates for measurement period indicated that relative to the post-instructions period, mean HR was significantly lower in the baseline, and significantly higher in the postural stressor period. Furthermore, changing the reference category revealed that mean HR was significantly lower in the baseline than in the rest period, indicating that HR had dropped slightly below resting levels after the postural stressor period. The estimates for week indicated that mean HR was significantly higher in week 2 than in week 3. The estimates for order indicated that relative to the fourth task, mean HR was significantly higher before the first and second task.

Table 3.4

GEE Analysis of Mean HR: Parameter Estimates

Source	*B*	Wald χ^2	Sig.	95% CI	
RP – PI	-0.56	1.44	.23	-1.48	0.36
PS – PI	10.97	259.92	< .001	9.64	12.31
BL – PI	-1.15	7.55	< .01	-1.96	-0.33
W 1 – W 3	0.99	0.36	.55	-2.24	4.22
W 2 – W 3	3.54	5.46	.02	0.57	6.52
TO 1 – TO 4	1.99	11.32	< .001	0.83	3.15
TO 2 – TO 4	1.40	10.63	< .01	0.56	2.23
TO 3 – TO 4	0.54	1.58	.21	-0.30	1.37
Intercept	75.36	1729.51	< .001	71.81	78.91

Note. Dependent variable: Mean HR. $N = 1335$. RP = Rest period. PI = Post-interventions. PS = Postural stressor. BL = Baseline. W = Week. TO = Task order.

The GEE analysis of mean HR change data found significant effects for the week (Wald $\chi^2 = 10.77$, $p < .01$), order (Wald $\chi^2 = 22.71$, $p < .001$), cardiovascular reference period (Wald $\chi^2 = 237.13$, $p < .001$), the week by order interaction (Wald $\chi^2 = 12.60$, $p = .05$), the week by reference period interaction effect (Wald $\chi^2 = 9.90$, $p < .01$), and the order by reference period interaction (Wald $\chi^2 = 10.31$, $p = .02$). Parameter estimates for this analysis are summarised in table 3.5. The parameter estimates for week indicated that HR change was not significantly different from week 3 in week 1, nor in week 2. The estimates for order indicated that mean HR change was significantly smaller before the first than before the fourth task. The estimate for cardiovascular reference period indicated that mean HR change was significantly greater in the postural than in the psychological reference period. The estimates for the week by order interaction effect indicated that relative to week 3, the difference in mean HR change between the first and the fourth task was significantly greater in week 1. The estimates for the week by reference period interaction effect indicated that relative to week 3, the difference in mean HR change between the postural and psychological reference period was significantly smaller in week 1 and in week 2. The estimates for the order by reference period interaction effect indicated that relative to the fourth task, the difference in mean HR change between the postural and psychological reference period was significantly smaller before the first task.

Table 3.5

GEE Analysis of HR Change: Parameter Estimates

Source	*B*	Wald χ^2	Sig.	95% CI	
W1 – W3	-0.86	0.44	.51	-3.38	1.67
W2 – W3	-1.85	2.22	.14	-4.27	0.58
TO 1 – TO 4	-2.29	3.96	.05	-4.55	-0.04
TO 2 – TO 4	-0.76	0.60	.44	-2.69	1.17
TO 3 – TO 4	-0.17	0.04	.84	-1.82	1.48
Postural – Psychological	13.64	93.87	< .001	10.88	16.40
$(W1 – W3)_{TO\ 1}$ – $(W1 – W3)_{TO\ 4}$	3.50	5.30	.02	0.52	6.48
$(W1 – W3)_{TO\ 2}$ – $(W1 – W3)_{TO\ 4}$	0.54	0.13	.72	-2.37	3.45
$(W1 – W3)_{TO\ 3}$ – $(W1 – W3)_{TO\ 4}$	0.87	0.33	.57	-2.10	3.84
$(W2 – W3)_{TO\ 1}$ – $(W2 – W3)_{TO\ 4}$	1.18	0.39	.53	-2.49	4.85
$(W2 – W3)_{TO\ 2}$ – $(W2 – W3)_{TO\ 4}$	1.24	0.72	.40	-1.63	4.12
$(W2 – W3)_{TO\ 3}$ – $(W2 – W3)_{TO\ 4}$	1.50	1.48	.22	-0.92	3.93
$(W1 – W3)_{Postural}$ – $(W1 – W3)_{Psychological}$	-3.96	9.45	< .01	-6.48	-1.44
$(W2 – W3)_{Postural}$ – $(W2 – W3)_{Psychological}$	-2.53	4.83	.03	-4.78	-0.27
$(TO\ 1 – TO\ 4)_{Postural}$ – $(TO\ 1 – TO\ 4)_{Psychological}$	-3.67	5.85	.02	-6.65	-0.70
$(TO\ 2 – TO\ 4)_{Postural}$ – $(TO\ 2 – TO\ 4)_{Psychological}$	-0.39	0.13	.72	-2.57	1.78
$(TO\ 3 – TO\ 4)_{Postural}$ – $(TO\ 3 – TO\ 4)_{Psychological}$	-0.51	0.14	.71	-3.15	2.13
Intercept	2.04	6.90	< .01	0.52	3.56

Note. Dependent variable: Mean HR change. TO = Task order. W = Week. *N* = 666.

3.4.3 GEE Analysis of Performance

The GEE analysis of performance is summarised in table 3.6. It found significant effects for week (Wald $\chi^2 = 41.25$, $p < .001$) and task (Wald $\chi^2 = 285.86$, $p < .001$). Parameter estimates indicated that relative to week 3, performance was significantly lower in weeks 1 and 2. Relative to the dart-throwing task, performance was significantly higher on the other three tasks. Neither cognitive, nor cardiovascular CAT were significantly associated with performance.

Table 3.6

GEE Analysis of Performance: Parameter Estimates

Source	*B*	Wald χ^2	Sig.	95% CI	
W 1 – W 3	-1.77	37.22	< .001	-2.34	-1.20
W 2 – W 3	-0.62	6.17	.01	-1.10	-0.13
SUT – DTT	8.88	155.64	< .001	7.48	10.27
NBT – DTT	9.89	244.26	< .001	8.65	11.13
BBT – DTT	2.25	28.75	< .001	1.42	3.07
TO 1 – TO 4	0.09	0.04	.84	-0.77	0.95
TO 2 – TO 4	0.55	1.12	.29	-0.47	1.56
TO 3 – TO 4	0.09	0.03	.86	-0.93	1.12
Cognitive CAT	0.15	2.15	.14	-0.05	0.35
Cardiovascular CAT	0.14	1.83	.18	-0.06	0.34
Intercept	6.50	112.92	< .001	5.30	7.70

Note. Dependent variable: Performance. $N = 328$. W = Week. SUT = Subtraction task. NBT = N-Back task. BBT = Bean-bag throwing task. DTT = Dart-throwing task. TO = Task order.

3.5 Discussion

The present study explored the variance components of cognitive and cardiovascular indicators of CAT states when examining variances between persons, tasks, time points, and their two-way interactions. It was hypothesised that at least one situational (state) component (task and/or time point), a person (trait) component, and at least one interaction component would be found (H1). This hypothesis was partially supported as person and interaction components were found for both cognitive and cardiovascular indicators of CAT, but there were no main effects for situational components on any variable. It also explored the variability of psychological (i.e., task engagement-related) cardiac reactivity relative to postural stressor-related cardiac reactivity (H2) and found that postural reactivity was more variable across repeated measurements than psychological reactivity. Finally, even though positive associations with performance were hypothesised (H3), neither cognitive, nor cardiovascular CAT were significantly associated with performance.

A variance components analysis provided new insights into the constituents of CAT states on the cognitive and cardiovascular level. In particular, the majority of the variance (59%) in cognitive CAT evaluations was explained by individual differences between persons (i.e., the person component) and interactions of individual differences with the tasks and time points (i.e., the person by task and person by week interaction components) at which the evaluations were reported. Differences between tasks also explained a considerable part of the variance (12%), but the task component did not reach statistical significance, potentially due to low statistical power. This pattern was found for the DRES and for both of its constituent variables (evaluations of perceived coping resources and situational demands). As such, these results suggest that individuals evaluate resources and demands in motivated performance situations in a dispositional fashion that is stable across tasks and time points. This is consistent with the findings of Lucas and colleagues (2012), who found that police officers are characterised by individual differences in appraising the stressfulness of work stressors. While there was no significant source of variation solely due to situational factors (e.g., task or time point) in the present study, there were significant person by task and person by week interaction components, suggesting that individuals' CAT evaluations may indeed be affected differently by different situations. Again, this is consistent with the findings of Lucas and colleagues (2012), who found that officers' individual difference characteristics interacted with stressor characteristics to explain some variance in stress appraisals. It is also consistent with findings from the social support literature that found social support (which might influence personal coping resources) to be largely determined by interaction components (Rees et al., 2012).

On the cardiovascular CAT index, a significant person component explained 16% of the variance. None of the other variance components were significantly different from zero, except for the error term, which explained 83% of the variance. Analysing the raw constituent variables of the cardiovascular CAT index (CO and TPR) as difference scores did show a slightly different picture. On CO change, there were significant person and person by week interaction components (jointly explaining 28% of the variance), indicating similar results as on the cognitive CAT evaluations. However, there were no person or person by week components on TPR change, which was characterised by a large error component (explaining 85% of the variance) and a significant week by task interaction that explained less than 1% of the variance. Hence, it appears that individual differences on the cardiovascular CAT index are largely due to stable cardiac, but not vascular reactivity profiles in motivated performance situations. However, the large error terms on the cardiovascular variables pose the question of whether a three-way interaction component (i.e., person by week by task) would have been significant if the study design had included a fourth factor. Since the study included only three factors, this question could not be answered, as the three-way interaction was confounded with the error term.

The present findings have theoretical and applied implications that could guide research and identify potential targets for interventions. First, they provided evidence relevant to the theoretical and previously rarely examined trait or state question. In particular, the findings supported the notion that CAT evaluations can partly be explained by a stable disposition to evaluate motivated performance situations more consistently with a challenge or a threat state (Tomaka et al., 2018). However, they also implied that such a disposition could critically interact with situational factors to predict CAT states, which would be consistent with findings from social support research (Rees et al., 2012). Although situational factors were found to determine CAT states when interacting with personal factors, this study did not provide any support for the idea that a situational component is a significant determinant of CAT states in itself. .

The main applied implication of the present study is that when testing potential challenge-promoting interventions, one should consider a multi-method approach. This way, one could remain flexible enough to help both those individuals who are generally threatened (reflecting a personal disposition), and those who experience a threat state only in certain situations (reflecting a person by situation interaction). Thus, sport psychologists should prioritise the development of interventions that can be tailored according to individual needs over one-size-fits-all approaches. Two examples for such a flexible intervention that could be adapted to context and stable individual needs would be self-talk and imagery (Hardy, 2006; Hardy, Oliver, & Tod, 2009; Williams & Cumming, 2012). Whereas sport psychologists may be primarily focused on optimising athlete performance, another applied implication relates to the person component on cardiovascular CAT states and their associated health outcomes. As a threat state has been associated with various adverse health effects (Blascovich, 2008b), and this study indicated stable tendencies in individuals' cardiovascular CAT responses, one might try to use cardiovascular CAT states to predict health outcomes. Thus, preventative medicine might potentially make applied use of CAT measurements to identify persons at high risk for health problems such as cardiovascular disease (Blascovich, 2008b).

This study also examined HR reactivity across reference periods (i.e., psychological/task engagement versus postural stressor), task order (i.e., each of the first three versus the last task), and time points (i.e., each of the first two versus the third). The results indicated that participants exhibited a cardiac response consistent with task engagement across tasks and weeks, and that a three-minute period after a postural stressor was sufficient to let HR return to resting values. Mean HR tended to be lower in the last, compared to the first, task and week. This finding might reflect lower general arousal levels in the later tasks/weeks as participants became habituated to the motivated performance situation (which did not change over time), although importantly this decrease in mean HR was not connected to an attenuation in reactivity. The results of Kelsey and colleagues (2004) were different from those of the present study, as they found a cardiovascular adaptation (i.e., decrease of cardiac reactivity) across tasks. However, there were some key differences between their and the present study that prevented a direct comparability with the present findings. For example, they did not compare baseline values against a mental preparation period, but against task performance. Also, they did not measure cardiovascular reactivity across different weeks or different tasks, as measurements were taken on only one day and one task (performed multiple times with intermittent baselines). They also did not compare psychological against postural reactivity as a control measure.

The question of whether this potential habituation affected task reactivity in the two reference periods (postural stressor versus psychological) differentially was answered by the analysis of mean HR change. Significant interaction effects between reference period and task order, as well as week, showed that contextual factors did change the difference between postural and psychological reactivity. Precisely, the difference between postural and psychological reactivity was smaller in weeks 1 and 2 than in week 3. Also, the same difference was smaller before the first than before the fourth task. Table 3.5 indicates that this is likely due to an increase in postural stressor reactivity from week 1 to week 3, as well as from the first to the fourth task, whereas psychological reactivity was relatively stable across measurements. This is inconsistent with the findings of Kelsey and colleagues (2004), which showed that psychological reactivity decreased across repeated measurements. However, the same caveats mentioned in the above paragraph also apply here, limiting the conclusions drawn from this comparison.

Neither cognitive, nor cardiovascular CAT were significantly related to performance across weeks and tasks, although the trends were of the predicted direction on both variables. The lack of a significant positive association with performance is inconsistent with the predictions of the BPSM and the findings of chapter 2, which found a challenge state to be superior to a threat state. Blascovich and Mendes (2000) highlighted that cognitive self-reports may be limited by cognitive distortions and low ability to accurately assess personal coping resources and situational demands. However, as the cardiovascular CAT variable used in this study avoids these limitations, one might still wonder whether other factors could explain the absence of a significant effect. As the variance components analysis found a large error component on the cardiovascular CAT variable, a hypothetical explanation might be the relatively low variation in CAT states throughout the study (i.e., between people and situations). In simple terms, this would imply that participants might not have experienced CAT states that were heterogeneous (i.e., extreme) enough to provoke meaningful performance differentials.

3.5.1 Limitations and Future Directions

The generalisability of the present findings may be limited by the low ecological validity of the study. As the testing environment was somewhat artificial, a real-world situation, such as a sport competition, might have provided greater ecological validity for the variance components inferred from the present data. The ecological validity of the motivated performance situation at hand is important because a highly ecologically valid study setting might provoke greater self-relevance, and thereby greater task engagement than the artificial competitive setting of the present study. Thus, there might be a potential for different magnitudes of the effects observed in the present study, although the general pattern of results should not differ. To increase ecological validity, a future study could collect repeated-measures data of CAT states in athletes at a series of competitions to examine whether this lack of ecological validity impacted the variance components in CAT states or the relationships between CAT states and performance. Although the present study was limited by low ecological validity, it nevertheless provided a sufficient motivated performance situation with its incentivised, pressurised performance context. Although potentially lower than in real-world competitions, cardiovascular task engagement was sufficient, as evidenced by significant HR increases in response to the task instructions across tasks and weeks.

The present findings suggest directions for future research. As CAT states appear to vary predominantly as a function of individual differences, experimental research could attempt to develop interventions that help those individuals with a general tendency to experience a threat state (e.g., psychoeducation or cognitive-behavioural interventions to improve dispositional self-efficacy or achievement goal orientation, physiological toughness-promoting interventions). Furthermore, the observed person by task and person by week interaction effects suggest that flexible interventions that can be adjusted to the context in which specific individuals experience a threat state or wish to intensify a challenge state (e.g., self-talk) might be most promising. Hence, future research should work toward a multi-method toolkit that can help both individuals who habitually experience a threat state, as well as individuals whose threat state experience is contingent on specific situational factors. For example, a long-term mindfulness or attentional training might prove more helpful for the former group, as they might have a generalised difficulty to get psychologically attuned to motivated performance situations. In contrast, the latter group might benefit more from specialised training, as their threat experience might be contingent on specific technical or psychological aspects of the task to be performed. The findings on HR reactivity indicated that future research could study fluctuations in postural HR reactivity in more detail, as it is unclear whether the observed increases were due to psychological (e.g., invested effort) or physiological factors (e.g., greater cardiac load with later tasks).

3.5.2 Conclusion

This study showed measured CAT states on the cognitive and cardiovascular level and found that CAT states largely vary between people, although some parts of CAT states also vary differentially between people across different tasks and time points.

This is consistent with prior research and presents important directions for future research toward challenge-promoting interventions, which should target person- and interaction-related sources of CAT states. On the contrary, situational factors did not emerge as significant variance components of CAT states when examined in isolation of personal factors. This study also showed that cardiac reactivity to psychological stressors is relatively stable across measurements, whereas postural stressor-related cardiac reactivity increased throughout tasks and weeks. This indicates that task engagement may not be an issue when conducting repeated-measures research on CAT states, although task engagement should continue to be monitored.

Table 3.7

Summary of Chapter 3 and Preview of Next Chapter

Chapter	Aim	Findings
3	To partition the variance in CAT states into personal (person), situational (task, week), and interaction components.	Significant person components were found for cognitive and cardiovascular CAT variables and explained 16-39% of the variance. Person by week and person by task interaction components were found on cognitive CAT variables only (jointly explaining 17-22% of the variance).
	Rationale for next chapter The study in chapter 3 was limited by low ecological validity (due to the laboratory-based testing setting with slightly artificial tasks). Thus the next study should examine the variance components of CAT states in athletes before real-world competitions.	
Chapter	Aim	Findings
4	To partition the variance in CAT states into personal (athlete) and dynamic [competition(athlete)] components.	

Chapter 4

A Repeated-Measures Examination of Challenge and Threat States in Competitive Trampoline Gymnastics

4.1 Abstract

A systematic review indicated that a challenge state relates to better performance than a threat state, and a first repeated-measures study has examined challenge and threat states in a laboratory context. However, no repeated-measures study has examined the relationship between challenge and threat states and performance in the field at elite sport competitions. This study examined the relationship between cognitive and cardiovascular indicators of challenge and threat states and performance; partitioned the variance in challenge and threat states into athlete, competition, and interaction components; and compared two different cardiovascular challenge and threat indices (based on silent imagination of task preparation versus a speech about task preparation). Thirty elite-level trampoline athletes (17 females, $M_{\text{Age}} = 14.6$ years, $SD = 3.4$) participated in three measurements taken before three out of six competitions, using a nested design. Cognitive evaluations consistent with a challenge state (personal resources matching or outweighing situational demands) were associated with worse performance than those consistent with a threat state ($B = -4.14$, $p < .01$), although age appeared to moderate this relationship. The effects of cardiovascular challenge and threat variables on performance were inconsistent, with the speech-based measure being a better predictor of performance than the silent imagination-based measure. The variance components analysis revealed significant interaction components between athlete and competition nested within measurement on all outcomes, explaining between 27.7% and 59.3% of the variance. These findings challenge the predictions of the biopsychosocial model in child and adolescent populations, and direct the development of interventions toward person-specific approaches.

4.2 Introduction

A number of sports require athletes to execute an extensively rehearsed routine of movements in a competition, for example trampoline gymnastics, high diving, and ski jumping. Coaches and applied sport psychologists may be concerned that the psychological pressure and stress associated with the competitive environment may provoke short-term and, in the worst case, long-term negative outcomes for their athletes (e.g., Hill, Cheesbrough, Gorczynski, & Matthews, 2019). Therefore, it is important to identify variables that predict performance under and capability to deal with psychological pressure before and in competitions. Cognitive and cardiovascular measures of CAT states are promising candidates for such variables, as chapters 1 and 2 have shown. However, limited research has examined their impact across multiple competitions and at an elite level, which is a gap in the literature that chapters 1 and 2 recommended to address. The primary purpose of this study thus was to measure CAT states at multiple elite-level trampoline competitions, to analyse personal and situational influences across measurements, and to predict performance with CAT states. At the same time, it also had the secondary purpose of comparing different ways of calculating a cardiovascular CAT measure.

As reviews have shown (see Behnke & Kaczmarek, 2018; chapter 2), many studies have supported the prediction of the BPSM that a challenge state relates to better performance than a threat state. Among other contexts, a challenge state has been associated with better performance than a threat state in baseball and softball (Blascovich et al., 2004), golf (Moore, Wilson, et al., 2013), netball, and cricket (Turner et al., 2012; Turner et al., 2013). These observational studies in the sport context have measured CAT states and predicted performance mostly on the same day, although one study predicted average performance throughout the competitive season, thereby demonstrating a considerable predictive validity of cardiovascular CAT measures (Blascovich et al., 2004).

A repeated-measures study in an elite sports context could also present a valuable extension of the findings on variance components of CAT states in chapter 3. As in chapter 3, a generalisability theoretical approach (Lakey, 2016; Shavelson & Webb, 2005) could be used if CAT states were measured in a group of athletes at several different competitions. Analogous to the person, situation, and person by situation interaction components in chapter 3, the variance in an applied sports context could be divided into components for the athlete, the competition, and the athlete by competition interaction. Chapter 3 revealed significant person (on cognitive and cardiovascular variables) and person by situation interaction components (on cognitive variables) in CAT states. However, the study in chapter 3 was limited by its artificial nature, for example due to testing university students and staff members in a controlled laboratory environment. Therefore, this study set out to replicate the findings from chapter 3 while avoiding the associated limitations. For this purpose, a group of elite level athletes (person component) was measured before different real-world competitions (situation component) to explain the variance in CAT states. Due to limited availability of athletes at competitions (not all athletes performed at all competitions), the study used a nested design wherein competition was nested within athlete (Lakey, 2016). In a naturalistic study setting such as this one, a nested design may be a helpful option for making the most of the collected data. However, the drawback of nested designs is that they confound the situation with the person by situation interaction component. For example, if athlete A performs at competitions 1, 2, and 3; and athlete B performs at competitions

1, 2, and 4; then only performance at competitions 1 and 2 can be analysed in a fully-crossed design where person, situation, and person by situation interaction components can be distinguished. A nested design is able to use data from all competitions, but this leads to the situation component (differences between competitions 1-4) being confounded with the person by situation interaction component (varying differences between athletes A and B across competitions 1-4). Thus, the present study partitioned the variance into a person component (athlete) and a dynamic component (competition nested within athlete). These components represented variance explained by differences between athletes (subsequently: “athlete”) and differences between competitions as nested within athletes [subsequently: “competition(athlete)”].

Another issue in CAT research that has not been previously addressed relates to the measurement of cardiovascular CAT responses. As mentioned above, cardiovascular CAT measurements typically involve a resting baseline period and a task-specific reactivity period. However, this task-specific reactivity period has differed in past research, with some studies having used cardiovascular data recorded during a speech about the respective task/sport (e.g., Blascovich et al., 2004) and others having used data recorded during a silent period during which participants imagined preparing for or performing the task (e.g., Moore et al., 2014). At this point, it should be noted that speech-based cardiovascular data may potentially be confounded by processes involved in speech production (e.g., muscular activity, changes in respiratory patterns, cognitive load). Neither the BPSM, nor previous empirical studies have examined differences between speech-based and silent imagination-based indicators of CAT states, and their relationship with performance. However, comparisons between studies indicate that the cardiac response associated with task engagement (i.e., HR reactivity) is greater in studies using speech-, rather than imagination-based variables. For example, in typical speech-based designs, HR reactivity ranged from 15.0-27.2 bpm (Blascovich et al., 2004; Mendes et al., 2002; Rith-Najarian et al., 2014). In typical imagination-based studies, HR reactivity ranged from 5.3-10.8 bpm (Moore, Wilson, et al., 2013; Moore et al., 2014; Vine et al., 2013). As these previous numbers are not perfectly commensurable due to emerging from different samples in different contexts, this study compared a cardiovascular CAT index based on silent imagination of task preparation with one that was based on speaking about the same imagination in the same participants. This was done to control for potential confounding speech production-related influences on HR. It also examined potential differences in how the two indices relate to performance.

The current study focused on performance in individual trampoline gymnastics. In this sport, athletes typically need to perform two qualifier routines composed of 10 jumps each. Each jump consists of a set of transversal and longitudinal body rotations that determine its difficulty (Fédération Internationale de Gymnastique, 2017). The first routine usually comprises mandatory jumps dictated by the organising committee, whereas the second routine comprises jumps freely chosen by the athlete and their coach. A panel of judges rates the difficulty and execution of each jump and calculates an overall performance score for each routine. Typically, the eight best performers in the two qualifier routines participate in a freely chosen final routine that determines the winner of the competition.

In sum, this study examined the repeated-measures relationship between CAT states (measured on the cognitive and cardiovascular level) and performance at several elite trampoline gymnastics competitions, as well as the variance components of CAT states in this setting. A secondary question of the study was whether a silent imagination-based and a speech-based cardiovascular CAT index differ regarding their HR reactivity and their relationships with performance. I hypothesised that a challenge state would relate to better performance than a threat state across competitions (H1). Consistent with the findings of chapter 3, I also hypothesised that a person component (athlete) would explain a significant percentage of the variance in cognitive evaluations and cardiovascular responses, and that a dynamic component (competition nested within athlete) would explain a significant percentage of the variance in cognitive evaluations (H2). Regarding potential differences between a silent imagination-based and a speech-based cardiovascular CAT index, I hypothesised greater HR reactivity in the speech-based index than in the silent imagination-based index (H3). I had no specific hypothesis regarding differential relationships of the cardiovascular CAT indices with performance, but explored potential differences.

4.3 Method

4.3.1 Participants

The sample consisted of 30 elite trampoline gymnasts (17 female, 13 male) from 10 different clubs spanning all five age categories (11-12, 13-14, 15-16, 17-21, and adult) competing on the national and international level. Age ranged from 10 to 22 years, with a mean of 14.6 years (*SD* = 3.4).

4.3.2 Materials

4.3.2.1 Cardiovascular data. The Portapres model-2 was used (for details, see chapter 3, p. 74).

4.3.2.2 Demand and resource evaluations. Four items assessed demand and resource evaluations: “How demanding do you expect the upcoming task to be?” and “How stressful do you expect the upcoming task to be?” for demands, and “How able are you to cope with the demands of the upcoming task?" and "How well do you think you can manage the demands imposed on you by this task?" for resources (Schneider, 2008). All items were scored on a seven-point Likert scale anchored by *not at all* (1) and *extremely* (7). A cognitive CAT variable (termed "Cognitive CAT 1") was created by subtracting the first demands item from the first resources item, meaning that possible scores ranged from -6 to 6 and denoted more challenge as values increased. For the variance components analyses reported in this chapter, a second cognitive CAT variable (termed "Cognitive CAT 2") was created by subtracting the second demands item from the second resources item.

4.3.3 Procedure

The study obtained institutional ethics approval. Before participating, each athlete provided written informed consent. In the case of underage athletes, written informed consent was obtained from both parents/caregivers. The study took place at various national and international trampoline jumping competitions and at the training sites of the participating trampoline clubs during the respective last training sessions before the competitions. The study comprised three measurement sessions, each of which consisted of a cardiovascular testing part and a subsequent questionnaire part. The measurements took place at different competitions for different athletes due to the competitive schedules and limited availability of most athletes. Thus, competitions were nested within measurement sessions. Figure 4.1 graphically represents this nested design. The six competitions and respective attendances were: 1) the last qualifier competition for the world championships/World Age Group Competitions[7] (30

[7] The World Age Group Competition is the equivalent of the world championships for the categories 11-12, 13-14, 15-16, and 17-21. It was held the week after the world championships in the same venue.

attending), 2) the world championships/World Age Group Competitions (18 attending), the national club championships (6 attending), as well as the first (18 attending), second (8 attending), and third qualifier competition for the national individual championships (10 attending).

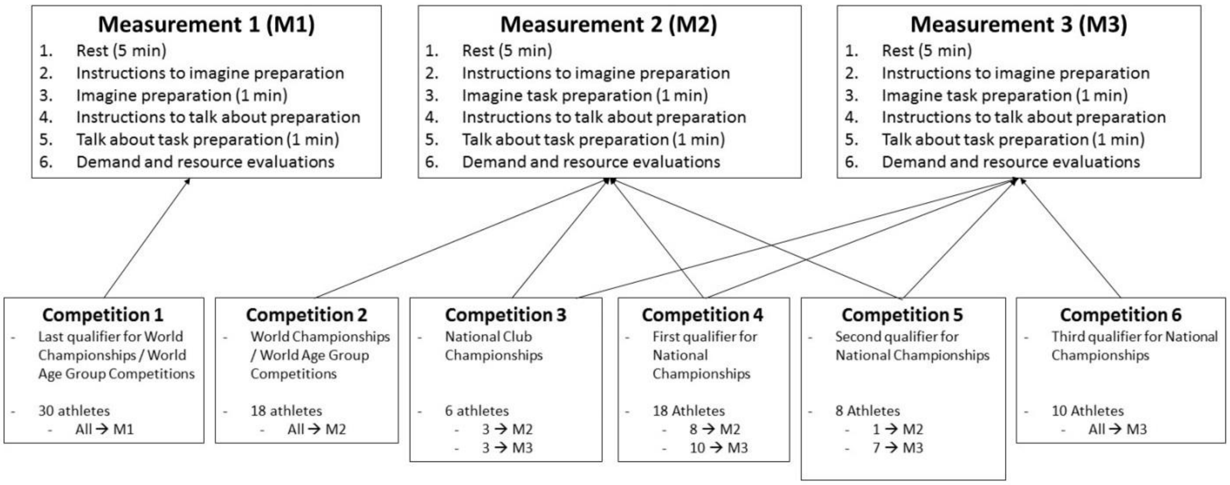

Figure 3.1. Overview of Nesting within Measurements.

For the cardiovascular testing session, the experimenter placed the Portapres cuff around the left ring finger of the athlete and placed the Portapres height correction sensor around the left arm at the height of the sternum. In case of signal problems, the middle or index finger was used instead. The cardiovascular testing period started with a five-minute baseline period during which the athlete was instructed to rest and relax. After the five-minute baseline period had elapsed, the experimenter went on to deliver the following instructions to the athlete:

> The rest period has now finished. We would now like to ask you to imagine your upcoming competition. Think of the last minute before starting your routine. This is the most important part of the experiment. While you imagine the preparation for your competition, we will record heart rate and blood pressure

> data for one minute. Please now think of the last minute before starting your competitive routine for one minute.

Cardiovascular data were recorded for one minute after these instructions to provide reactivity data for the first cardiovascular CAT index (silent imagination-based; subsequently termed "Cardiovascular CAT 1"). After the minute had elapsed, the experimenter gave the following instructions to the athlete:

> For the next one minute, we would like you to describe out loud your feelings and thoughts that you are going to have during the last minute before starting your routine. We will again record heart rate and blood pressure data during the next minute. Then, we would ask you to fill in the questionnaire items. After that, you will be done for the day.

As the athlete talked, cardiovascular data were recorded for another minute to provide reactivity data for the second cardiovascular CAT index (speech-based; subsequently termed "Cardiovascular CAT 2"). After this minute was recorded, the experimenter announced that the cardiovascular data collection was complete, removed the Portapres, asked the athlete to complete the self-report measure of demand and resource evaluations of their upcoming competition, and thanked them for their participation. Performance scores for each routine at the competition were retrieved from the official results publication of the respective competition. Because most athletes only competed in the two qualifier routines, the analyses did not include final routine data.

4.3.4 Statistical Analysis

Consistent with previous research using the BPSM of CAT (e.g., Mendes et al., 2007), mean HR, TPR, and CO values were calculated for the final baseline minute, the minute after the first set of instructions, and the minute after the second set of instructions. Seven univariate outliers (values more extreme than three standard deviations from the sample mean at the respective time point; Stevens, 2009) were winsorised to be 1% more extreme than the next non-outlying score (as Shimizu et al., 2011). The baseline values for CO and TPR were then regressed on their respective reactivity values with the standardised residuals being saved to create residualised change scores in order to adjust for baseline differences (Burt & Obradovic, 2013). The TPR residualised change scores were then subtracted from the CO residualised change scores to create a single cardiovascular CAT index (i.e., cardiovascular CAT 1 / cardiovascular CAT 2) which is common in research employing a BPSM framework (e.g., Vine et al., 2013).

To test task engagement, a GEE analysis predicted mean HR with measurement period (the final baseline minute being the reference category) and measurement session (the first measurement session being the reference category), both of which were specified as within-subjects factors. For task engagement, the difference between the final baseline minute and each of the other two measurement periods was examined. To test the difference between the speech- and imagination-based measurement periods (H3), the analysis was repeated with the speech-based minute selected as the reference category.

To test H1, two GEE analyses were conducted to predict the respective performance scores (routines 1 and 2) with cognitive CAT 1, cardiovascular CAT 1, cardiovascular CAT 2, age, sex, and the respective interaction effects of the CAT variables with age and sex (i.e., Cognitive CAT 1*Age, Cognitive CAT 1*Sex, Cardiovascular CAT 1*Age, Cardiovascular CAT 1*Sex, Cardiovascular CAT 2*Age, Cardiovascular CAT 2*Sex.). The GEE models were used because a GEE analysis enables the test of relationships between a set of categorical and continuous independent variables (including their interactions) and a dependent variable across different time points, which is a parsimonious alternative to conducting separate analyses at each time point. All GEE analyses assumed an independent correlation structure and excluded cases with missing data. All analyses used a significance level of $\alpha = .05$.

To test H2, six variance components analyses analysed the following outcomes: cognitive CAT, resource evaluations, demand evaluations, cardiovascular CAT, raw CO change, and raw TPR change. Using the restricted maximum likelihood estimation method, each variance components analysis partitioned the total variance of the outcome into athlete, item, competition nested within athlete [subsequently: “competition(athlete)”], athlete by item interaction, and error components. The item component was added because in generalisability theory, the highest order interaction is confounded with the error term (Lakey, 2016) and nested effects statistically count as interaction effects. Thus, an additional component was needed to yield a meaningful competition(athlete)] interaction effect, because it would otherwise have been confounded with the error term. The addition of an item component is commonplace in generalisability theory research (e.g., Lakey et al., 2004). The item component was added by treating cognitive CAT 1 (same for its constituent demand and resource evaluation items) as “item 1” and cognitive CAT 2 (and its constituent items) as “item 2”. Likewise, cardiovascular CAT 1 (and the respective raw CO and TPR change variables) was treated as “item 1” and cardiovascular CAT 2 (and the respective raw CO and TPR change variables) as “item 2”. Components were tested for significance by computing a 95% confidence interval and examining whether it excluded zero (as Lakey et al., 2004). To do this, the following formula was used, where *x* denotes the respective variance component and *var* denotes the variance of the respective variance component:

$$95\%\ \mathrm{CI}(x \pm \sqrt{var_x} \times 1.96)$$

4.4 Results

Table 4.1 reports descriptive statistics for key variables by measurement session. There was one case of missing performance data as one athlete did not compete at the competition following their third measurement session. There were two cases of missing cognitive data (two athletes did not report any cognitive evaluations at their second measurement session). There were eight cases of missing cardiovascular data due to equipment problems (five in the first, one in the second, and two in the third measurement session). Missing data were excluded pairwise.

Table 4.1

Descriptive Statistics by Measurement Session

	Measurement 1		Measurement 2		Measurement 3	
	M	*SD*	*M*	*SD*	*M*	*SD*
1. Performance – Routine 1	40.50	4.22	39.92	5.47	41.41	3.16
2. Performance – Routine 2	41.06	11.95	42.04	12.80	41.21	12.94
3. Cognitive CAT 1	0.57	1.76	0.46	1.50	1.03	1.99
4. Cardiovascular CAT 1	0.00	1.69	0.00	1.82	0.00	1.78
5. Cardiovascular CAT 2	0.00	1.84	0.00	1.85	0.00	1.84
6. Task Engagement 1	2.91	6.10	3.50	5.24	1.96	4.73
7. Task Engagement 2	6.47	6.66	8.69	5.23	7.39	8.05
8. CO Reactivity 1	0.15	0.77	0.28	0.52	0.00	0.91
9. CO Reactivity 2	0.22	0.88	0.33	0.72	-0.16	1.06

10. TPR Reactivity 1	-0.02	0.54	0.00	0.15	0.06	0.24
11. TPR Reactivity 2	0.10	0.39	0.05	0.23	0.06	0.33

Note. Significance denoted by † $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

The GEE analysis of mean HR found a significant effect for measurement period (Wald $\chi^2 = 90.53$, $p < .001$). Parameter estimates showed that HR increased significantly from baseline to post-instructions for both reactivity periods, thereby indicating sufficient task engagement [First reactivity minute: $B = 2.79$, 95% CI (1.42, 4.17), Wald $\chi^2 = 15.86$, $p < .001$; Second reactivity minute: $B = 7.57$, 95% CI (6.01, 9.13), Wald $\chi^2 = 90.14$, $p < .001$]. Repeating the analysis with the second reactivity minute as the reference category found that task engagement (i.e., raw HR reactivity) was significantly lower in the first, compared to the second reactivity minute [$B = -4.78$, 95% CI (-6.29, -3.26), Wald $\chi^2 = 38.36$, $p < .001$].

4.4.1 CAT and Competition Performance

The GEE analysis of performance (routine 1) found significant main effects for cognitive CAT 1 (Wald $\chi^2 = 11.03$, $p < .001$), cardiovascular CAT 2 (Wald $\chi^2 = 4.23$, $p = .04$), as well as significant interaction effects for cognitive CAT 1 by age (Wald $\chi^2 = 16.99$, $p < .001$) and cardiovascular CAT 2 by age (Wald $\chi^2 = 4.04$, $p = .04$). Table 4.2 presents parameter estimates for this analysis. The parameter estimate for cognitive CAT 1 indicated that cognitive evaluations consistent with a challenge state were associated with significantly worse performance than those consistent with a threat state [$B = -3.65$, Wald $\chi^2 = 11.33$, $p < .001$, 95% CI (-5.78, -1.53)]. The parameter estimate for cardiovascular CAT 2 indicated that speech-based cardiovascular responses consistent with a challenge state were associated with significantly better performance than those consistent with a threat state [$B = 3.83$, Wald $\chi^2 = 3.97$, $p = .05$, 95% CI (0.06, 7.60)].

The estimate for the significant cognitive CAT 1 by age interaction effect indicated a positive relationship, which can be interpreted as cognitive evaluations being more positively related to performance as age increased [B = 0.28, Wald χ^2 =16.99, p < .001, 95% CI (0.15, 0.42)]. The estimate for the significant cardiovascular CAT 2 by age interaction effect indicated a negative relationship, which can be interpreted as speech-based cardiovascular CAT responses being more negatively related to performance as age increased [B = -0.29, Wald χ^2 = 4.04, p = .04, 95% CI (-0.58, -0.01)].

Table 4.2

GEE Parameter Estimates for Routine 1

Source	*B*	Wald χ^2	Sig.	95% CI	
Cognitive CAT 1	-3.65	11.33	< .001	-5.78	-1.53
Cardiovascular CAT 1	-3.36	1.36	.24	-9.01	2.29
Cardiovascular CAT 2	3.83	3.97	.05	0.06	7.60
Sex: Male – Female	1.48	2.27	.13	-0.45	3.40
Age	0.33	2.63	.10	-0.07	0.72
Cognitive CAT 1_{Male} – Cognitive CAT 1_{Female}	-0.32	0.33	.57	-1.40	0.77
Cardiovascular CAT 1_{Male} – Cardiovascular CAT 1_{Female}	-0.91	1.91	.17	-2.19	0.38
Cardiovascular CAT 2_{Male} – Cardiovascular CAT 2_{Female}	0.89	2.92	.09	-0.13	1.90
Cognitive CAT 1 * Age	0.28	16.99	< .001	0.15	0.42
Cardiovascular CAT 1 * Age	0.27	1.51	.22	-0.16	0.70
Cardiovascular CAT 2 * Age	-0.29	4.04	.04	-0.58	-0.01
Intercept	35.06	173.21	< .001	29.84	40.28

Note. Dependent variable: Performance (routine 1). N = 79.

The GEE analysis of performance (routine 2) found significant interaction effects for sex by cardiovascular CAT 2 (Wald χ^2 = 5.02, p = .03) and cognitive CAT 1 by age

(Wald χ^2 = 4.96, *p* = .03). It also found marginally significant main effect trends for sex (Wald χ^2 = 3.77, *p* = .05), cognitive CAT 1 (Wald χ^2 = 3.38, *p* = .07), and cardiovascular CAT 1 (Wald χ^2 = 3.06, *p* = .08). Table 4.3 presents parameter estimates for this analysis. The parameter estimate for sex by cardiovascular CAT 2 indicated that the relationship between cardiovascular CAT 2 and performance was significantly less positive for male than for female athletes [*B* = -5.03, Wald χ^2 = 5.02, *p* = .03, 95% CI (-9.43, -0.63)]. The estimate for the cognitive CAT 1 by age interaction effect indicated that cognitive evaluations were more positively related to performance as age increased [*B* = 0.36, Wald χ^2 = 4.96, *p* = .03, 95% CI (0.04, 0.67)]. Estimates for the sex trend indicated that male athletes performed worse than female athletes [*B* = -6.07, Wald χ^2 = 3.77, *p* = .05, 95% CI (-12.19, 0.06)]. The trend for cognitive CAT 1 indicated that cognitive evaluations consistent with a challenge state were related to worse performance than those consistent with a threat state [*B* = -4.89, Wald χ^2 = 3.58, *p* = .06, 95% CI (-9.96, 0.18)]. The trend for cardiovascular CAT 1 indicated that imagination-based cardiovascular responses consistent with a challenge state were related to worse performance than those consistent with a threat state [*B* = -8.03, Wald χ^2 = 3.06, *p* = .08, 95% CI (-17.03, 0.96)].

Table 4.3

GEE Parameter Estimates for Routine 2

Source	*B*	Wald χ^2	Sig.	95% CI	
Cognitive CAT 1	-4.89	3.58	.06	-9.96	0.18
Cardiovascular CAT 1	-8.03	3.06	.08	-17.03	0.96
Cardiovascular CAT 2	-2.13	0.13	.72	-13.69	9.42
Sex: Male – Female	-6.07	3.77	.05	-12.19	0.06
Age	-0.86	2.72	.10	-1.88	0.16
Cognitive CAT 1_{Male} – Cognitive CAT 1_{Female}	0.80	0.28	.60	-2.15	3.75

Cardiovascular CAT 1_{Male} – Cardiovascular CAT 1_{Female}	1.08	0.28	.60	-2.94	5.11
Cardiovascular CAT 2_{Male} – Cardiovascular CAT 2_{Female}	-5.03	5.02	.03	-9.43	-0.63
Cognitive CAT 1 * Age	0.36	4.96	.03	0.04	0.67
Cardiovascular CAT 1 * Age	0.45	2.11	.15	-0.16	1.06
Cardiovascular CAT 2 * Age	0.27	0.35	.55	-0.62	1.16
Intercept	53.74	65.93	< 0.001	40.77	66.71

Note. Dependent variable: Performance (routine 2). *N* = 79.

4.4.2 Variance Components Analyses

The variance components analyses of cognitive CAT evaluations are detailed in Table 4.4. For cognitive CAT, the competition(athlete) component (explaining 39.2% of the variance) and the error component (explaining 38.0% of the variance) were significant. For resource evaluations, the athlete component (explaining 23.1% of the variance), the competition(athlete) component (explaining 25.3% of the variance), and the error component were (explaining 41.2% of the variance) significant. For demand evaluations, the competition(athlete) component (explaining 38.2% of the variance) and the error component were significant (explaining 39.1% of the variance).

Table 4.4

Variance Components Analyses of Cognitive Evaluations

Source	Component	Percentage of Variance	95% CI
Cognitive CAT			
Athlete	0.57	16.79	(-0.19, 1.32)
Item	0.00	0.00	N/A
Competition(Athlete)	1.33	39.20	(0.57, 2.08)*
Athlete*Item	0.21	6.07	(-0.16, 0.57)
Error	1.28	37.95	(0.82, 1.75)*
Resources			
Athlete	0.35	23.05	(0.01, 0.70)*
Item	0.09	5.86	(-0.19, 0.36)
Competition(Athlete)	0.39	25.26	(0.11, 0.67)*
Athlete*Item	0.07	4.58	(-0.10, 0.24)
Error	0.63	41.24	(0.40, 0.86)*

Demands			
Athlete	0.19	10.53	(-0.18, 0.57)
Item	0.06	3.02	(-0.14, 0.25)
Competition(Athlete)	0.71	38.24	(0.30, 1.12)*
Athlete*Item	0.17	9.11	(-0.06, 0.40)
Error	0.72	39.10	(0.46, 0.99)*

Note. Significant variance components are denoted by an asterisk (*). $N = 176$.

The variance components analyses of cardiovascular CAT responses are detailed in Table 4.5. For the cardiovascular CAT index, there were significant variance components for the competition(athlete) (explaining 59.3% of the variance) and the error term (explaining 26.8% of the variance). On raw CO change, there were significant components for the competition(athlete) (explaining 58.7% of the variance) and the error term (explaining 25.6% of the variance). On raw TPR change, there were significant components for the competition(athlete) (explaining 33.4% of the variance) and the error term (explaining 65.5% of the variance).

Table 4.5

Variance Components Analyses of Cardiovascular Variables

Source	Component	Percentage of Variance	95% CI
Cardiovascular CAT			
Athlete	0.30	9.30	(-0.42, 1.01)
Item	0.00	0.00	N/A
Competition(Athlete)	1.90	59.27	(0.99, 2.80)*
Athlete*Item	0.15	4.65	(-0.11, 0.41)
Error	0.86	26.78	(0.53, 1.18)*
CO Change (raw)			
Athlete	0.04	5.33	(-0.11, 0.19)
Item	0.00	0.00	N/A
Competition(Athlete)	0.41	58.73	(0.21, 0.60)*
Athlete*Item	0.07	10.35	(0.00, 0.15)
Error	0.18	25.58	(0.11, 0.24)*
TPR Change (raw)			
Athlete	0.00	0.00	N/A
Item	0.00	0.64	(0.00, 0.01)
Competition(Athlete)	0.04	33.44	(0.01, 0.06)*
Athlete*Item	0.00	0.46	(-0.01, 0.02)

Error	0.07	65.46	(0.04, 0.10)*

Note. Significant variance components are denoted by an asterisk (*). $N = 164$.

4.5 Discussion

The present study tested the hypothesis that a challenge state on the cognitive and cardiovascular level would be associated with better performance than a threat state across three competitions for elite-level trampoline gymnasts (H1), that a person component (athlete) would explain a significant percentage of the variance in all CAT measures, that a dynamic component [competition(athlete)] would explain a significant percentage of the variance in cognitive evaluations (H2), and that a speech-based cardiovascular CAT index would be associated with greater HR reactivity than a silent imagination-based index (H3). We also explored the potentially differential associations of the silent imagination-based and speech-based CAT indices with performance. There was mixed support for H1, as the speech-based cardiovascular CAT measure was the only CAT variable to be associated with performance in the predicted direction (for routine 1 only), whereas cognitive CAT evaluations were related to performance in the opposite direction. However, significant interactions of age with cognitive evaluations indicated that the latter finding may have been due to the young age of the sample. H2 was partially supported, as significant dynamic components were found throughout, but only one person component was found across the six CAT outcomes. H3 was supported, as HR reactivity was greater in the speech-based than in the imagination-based cardiovascular measurement period. Furthermore, the two cardiovascular CAT indices exhibited differences in how they related to performance, where the speech-based index was more positively related to performance than the imagination-based index.

Cognitive evaluations consistent with a challenge state (i.e., coping resources matching or exceeding situational demands) related to significantly worse routine 1 performance than those consistent with a threat state. Further, the same trend approached significance on routine 2 performance. This is inconsistent with the findings of Moore and colleagues, who found that cognitive evaluations predicted competitive performance in golfers (Moore, Wilson, et al., 2013). It is also inconsistent with a systematic review that found 76% of associations between cognitive CAT evaluations and performance indicated the superiority of a challenge over a threat evaluation (chapter 2). One potential explanation for these divergent findings could be the low age of the present sample, as young athletes may not yet be able to accurately assess their resources and competitive demands. Indeed, the cognitive evaluations by age interaction effects observed in the present study indicated that the relationship between cognitive evaluations and performance on both routines became more consistent with the BPSM and previous research findings as age increased. This idea was also supported by a previous study with a similarly young sample, which also failed to replicate the association between CAT evaluations and performance as predicted by the BPSM (Rith-Najarian et al., 2014). Although it does not explicitly list young age as a potential source of bias, the BPSM also acknowledges that compared to cardiovascular measures, self-report measures of CAT states have a higher risk of bias due to inaccurate assessments or little conscious awareness of psychological processes involving demand and resource evaluations, as well as other issues like self-presentation concerns (Blascovich & Mendes, 2000). Experience in the motivated performance task at hand, which might correlate with age in athlete samples) might be another factor to consider when examining potential moderators of the relationship between cognitive CAT evaluations and performance.

There was mixed evidence for the cardiovascular CAT variables. There was better routine 1 performance in athletes with a cardiovascular CAT 2 (i.e., speech-based) score more consistent with a challenge (relative to a threat) state. However, the other three associations were inconsistent as no significant relationships were observed, and cardiovascular CAT 1 approached significance in the opposite direction. Whereas the former finding was consistent with prior research using a speech-based cardiovascular CAT index (e.g., Blascovich et al., 2004) and supports the general pattern of results in the literature (see Behnke & Kaczmarek, 2018; chapter 2), the latter finding was inconsistent with the typical finding that a challenge state was superior to a threat state (e.g., Behnke & Kaczmarek, 2018; chapter 2). Unlike the negative association for cognitive CAT evaluations, which was moderated by age, it is unlikely that age played a significant role in the cardiovascular CAT findings, as the only significant interaction with age involved speech-based cardiovascular responses being *less* positively related to routine 1 performance with increasing age. However, it is noteworthy that another study with a comparably young sample also failed to replicate the association between cardiovascular CAT states and performance (mean age of 14.7 years; Rith-Najarian et al., 2014). On the other hand, Turner and colleagues (2013; mean age of 16.5 years) replicated the association between cardiovascular CAT and performance in an athlete sample only two years older than the present one. To further add to the ambiguity, Rith-Najarian and colleagues (2014) did not observe the relationship between cardiovascular CAT and performance using a speech-based CAT index (related to routine 1 performance in this study), whereas Turner and colleagues (2013) did observe the relationship using an imagination-based CAT index (not positively related to performance in this study). Thus, the present findings raise the question of why there were inconsistencies with the majority of previous CAT research and theory in terms of the relationship between cardiovascular responses and performance. Although there is no clear reason why CAT states should not generalise to children and adolescents, the BPSM has not clearly specified whether or not CAT states exist in children. This might be due to the relatively large variation in developmental status among individuals of the same chronological age in adolescence and childhood, which might make it more difficult to study the phenomenon of CAT states in these age groups.

The variance components analyses in the present study yielded results that were consistent between the cognitive and cardiovascular CAT outcomes, but were only partially consistent with the results of chapter 3. Precisely, significant dynamic [competition(athlete)] components were found on all outcomes, which in a nested design may represent both situational and person by situation interactional components. This implies that CAT states may change from one competition to another; change with the competition for some, but not all athletes; or that CAT states change from one competition to the next in all athletes, but in different directions or magnitudes. Certainly, the present finding highlights the need for a person-specific approach to CAT monitoring and interventions in elite sport, where interventions are selected based on whether they suit the personal profile of the recipient and how this profile interacts with the environment in motivated performance situations. The present finding partly builds on the results of chapter 3, which found person by situation interaction components on cognitive CAT evaluations, but not on cardiovascular responses. It is noteworthy that the percentages of variance explained by these components were larger in this study (25-59%) than in chapter 3 (17-22%). On average, they resembled those of Lucas and colleagues (2012), who found significant person by situation interactions in their data (explaining 38-41% of the variance), albeit on cognitive appraisals (not CAT evaluations) as outcomes. Despite the different outcome measure, a hypothetical explanation for the similarity between Lucas and colleagues' and the present findings could be that participants (police officers and athletes, respectively) rated stress responses to real-world scenarios (work stressors and competitions, respectively) rather than the laboratory-based motivated performance situations in chapter 3. Unlike chapter 3 and previous cognitive appraisal work (Lucas et al., 2012), this study found only one significant person component (on resource evaluations). It is unclear if this was due to the different sample, the nested study design, or another factor.

This study provided novel insight into how a cardiovascular CAT index based on silent imagination of the last minute before starting one's competition may differ from an index based on talking about the same last minute before the competition. Previously, Blascovich and colleagues (2004) had shown that a CAT index based on a speech about one's sport exhibited a different relationship with season performance in athletes than a CAT index based on a speech about friends. However, to my knowledge, no previous study has compared a CAT index based on silently imagining competition preparation with a CAT index based on talking about the same preparation to account for potential confounding influences of speech production on cardiovascular CAT responses. Neither of the two indices predicted routine 2 performance, but differential relationships with routine 1 performance were found. In particular, the association for the speech-based, but not the imagination-based CAT index exhibited a positive trend on routine 1 performance (i.e., consistent with the BPSM). A potential explanation for this finding could be that talking about competition preparation is more engaging than imagining competition preparation, as talking to an experimenter might require more vivid imagery to come up with more concrete and detailed descriptions of the preparation. Such an increase in task engagement might in turn produce a stronger relationship between cardiovascular indicators of CAT and performance. This idea would be supported by the higher HR reactivity in the speech-based reactivity period. Thus, it appears that the metabolic demands of speech production did not confound the validity of the cardiovascular CAT index, although they may indeed have provoked greater HR reactivity.

An ancillary analysis also showed that the speech-based CAT index featured greater HR reactivity than the silent imagination-based index. This is consistent with the results of previous research (e.g., Blascovich et al., 2004; Moore, Wilson, et al., 2013; Moore et al., 2014; Seery et al., 2010), although previous data only allowed for cross-comparisons between studies and therefore may have been confounded by factors other than the silent imagination-speech distinction. The current study thus presents the first within-subjects comparison confirming greater speech-based than silent imagination-based HR reactivity, having controlled for contextual factors (e.g., outcome task, participants, incentives). However, it is unclear whether the greater HR reactivity during the speech reflects greater task engagement, greater cardiac activity due to the physiological demands of speech production, or a cumulative effect of both the imagination of and the speech about competition preparation. Other studies have also recorded cardiovascular data during task performance (e.g., Scholl et al., 2017), which would have been interesting, but was not feasible in this study as the outcome task prevented reliable cardiovascular measurements during performance.

Some limitations of this study should be noted. The naturalistic study design prevented sampling more participants to balance data lost to dropout. The fact that the sample contained mostly adolescent participants is not a limitation in itself, but the small number of adult athletes in the sample prevented robust conclusions about how CAT states and their relationships with performance change with age. The Portapres model-2 did not allow for the calculation of VC and examining a task engagement index based on both HR and VC might have been more robust than HR reactivity only (e.g., Streamer et al., 2017). Finally, the comparison between imagination- and speech-based cardiovascular data could have benefited from a stronger design. For example, the order of the two reactivity minutes could have been counter-balanced and a second rest period could have been added between the two reactivity minutes to let values return to baseline after the first reactivity minute.

The present findings highlight a need for future research to examine age as a potential moderator of the relationship between CAT states and performance. Depending on the results, the BPSM might need to re-specify its predictions as not applicable or applicable only under specific conditions to the study of children and adolescents. The present study could also be repeated with a sample of more homogenous age to examine the undistorted relationships between CAT states and elite-level trampoline performance. Finally, a study using a fully-crossed (i.e., non-nested) design where all athletes experience the same competitions might be needed to examine whether differences between the results of this chapter and chapter 3 were due to the different sample and context, or due to the different study designs (i.e., fully-crossed versus nested).

4.5.1 Conclusion

This study was the first to examine the relationship between cognitive and cardiovascular CAT measures across multiple competitions throughout the season of elite-level athletes. Contrary to expectations, cognitive CAT evaluations were negatively related to performance, but an interaction effect with age indicated that this may have been due to the young age of participants. Cardiovascular CAT did not consistently relate to performance, but a speech-based variable predicted routine 1 performance in the predicted direction. Variance components analyses showed that cognitive, as well as cardiovascular indicators of CAT states varied largely as a function of competitions nested within athletes. This study was also the first to provide a within-subjects comparison of a silent imagination-based and a speech-based cardiovascular CAT index. This comparison showed that the speech-based CAT index had a more positive relationship with performance and greater HR reactivity than the silent imagination-based index, implying that a speech-based index may be more useful in predicting elite-level competitive performance. The present findings provide valuable information for sport professionals and CAT researchers alike as they provide a first repeated-measures examination of CAT states in elite athletes and highlight a potential need to specify new boundary conditions to the BPSM. Thus, future research is encouraged to examine age as a moderator of the CAT-performance relationship.

Table 4.6

Summary of Chapter 4 and Preview of Next Chapter

Chapter	Aim	Findings
4	To partition the variance in CAT states into personal (athlete) and dynamic [competition(athlete)] components.	Significant competition(athlete) components were consistently found and explained 25-59% of the variance in CAT states.
	Rationale for next chapter The finding that CAT states significantly varied as a function of personal or person by situation interactional factors indicated that interventions aiming to optimise CAT states and thereby improve performance could target processes linked to these factors. Self-talk is one example of a process that occurs organically (i.e., autonomously) and varies between people (and potentially also as a function of person by situation interactions). However, it can also be strategically used as an intervention to improve performance. Thus, chapter 5 examined whether two strategic self-talk interventions could impact CAT states in a motivated performance situation (relative to control self-talk).	
Chapter	Aim	Findings
5	To examine whether instructional and motivational self-talk promote a challenge state.	

Chapter 5

The Influence of Self-Talk on Challenge and Threat States and Performance

Published as:

Hase, A., Hood, J., Moore, L. J., & Freeman, P. (2019). The influence of self-talk on challenge and threat states and performance. *Psychology of Sport and Exercise*, *45*, 101550. doi: 10.1016/j.psychsport.2019.101550

5.1 Abstract

5.1.1 Objectives

A challenge state has been consistently associated with better performance than a threat state. However, to date, challenge-promoting interventions have rarely been tested. Therefore, this study investigated whether instructional and motivational self-talk promoted a challenge state and improved task performance.

5.1.2 Design

A three-group, randomised-controlled experimental design was used.

5.1.3 Method

Sixty-two participants (52 males, 10 females; M_{age} = 24 years, *SD* = 6) were randomly assigned to one of three self-talk groups: instructional, motivational, or control. Participants performed four dart-throwing tasks. Cognitive and cardiovascular measures of challenge and threat states were recorded before the first and final task.

5.1.4 Results

The motivational, but not the instructional group, improved their performance between the first and final tasks more than the control group. Self-talk had no effect on the cognitive or cardiovascular challenge and threat measures. However, evaluating the task as more of a challenge (i.e., coping resources match/exceed task demands) was related to better performance. Cardiovascular reactivity more reflective of a challenge state (i.e., higher cardiac output and/or lower total peripheral resistance reactivity) was more positively related to performance in the motivational than in the control group, and in the control than in the instructional group.

5.1.5 Conclusions

Motivational self-talk improved performance more than control self-talk. Furthermore, motivational self-talk may have strengthened, whereas instructional self-talk may have weakened, the relationship between challenge and threat states and performance. Hence, athletes in a challenge state may benefit from motivational self-talk, whereas those in a threat state may profit from instructional self-talk.

5.2 Introduction

In elite sport, it is common to see some athletes choke, whereas others excel under pressure. The biopsychosocial model (BPSM) of challenge and threat states (Blascovich, 2008), and the Theory of Challenge and Threat States in Athletes (TCTSA; Jones, Meijen, McCarthy, & Sheffield, 2009) provide explanations for such instances of performance variability. The theories conceptualise challenge and threat (CAT) states as distinct patterns of cognitive evaluations and physiological responses in motivated performance situations. It is noteworthy that a seemingly unrelated theory that provides a framework for the study and application of self-talk within sport (Hardy, Oliver, & Tod, 2009), describes self-talk as a phenomenon with considerably consistent effects. Thus, this study tested whether self-talk, a widely researched phenomenon in sport, influenced CAT states.

Motivated performance situations (e.g., sporting competitions, university exams, job interviews) are characterised by their potentially stressful nature, and require an active coping effort or an instrumental cognitive and/or behavioural response, to attain an important and self-relevant goal (Blascovich, 2008). In these situations, CAT states occur on a single bipolar continuum, which can be described in terms of underlying cognitive evaluations and accompanying physiological responses (Blascovich, 2008). Due to the continuous nature of CAT states, relative rather than absolute differences in CAT are often examined. Toward the challenge end of the continuum, athletes evaluate that their coping resources match or exceed situational demands. Toward the threat end, athletes evaluate that coping resources fall short of situational demands. It should be noted that these evaluations are subjective rather than objective. The BPSM posits that the balance of evaluated coping resources to situational demands engenders specific physiological responses. Both CAT states require task engagement, which is marked by increases in heart rate (HR; number of heart beats per minute) and ventricular contractility (VC; contractile state of the left ventricle). A challenge evaluation, however, is associated with a cardiovascular reactivity pattern consisting of relatively greater cardiac output (CO; volume of blood ejected by the left ventricle per minute) and lower total peripheral resistance (TPR; degree of systemic peripheral vascular constriction), whereas a threat evaluation is linked to a pattern composed of relatively lower CO and greater TPR (Tomaka, Blascovich, Kelsey, & Leitten, 1993).

Both the BPSM and TCTSA specify that a challenge state is related to better performance than a threat state (Blascovich, 2008; Jones et al., 2009). Although a recent meta-analysis noted that the effect may be small (Behnke & Kaczmarek, 2018), a challenge state has been associated with superior performance relative to a threat state in 74% of studies conducted across various tasks and contexts (e.g., baseball and softball, golf putting, laparoscopic surgery; see Hase, O'Brien, Moore, & Freeman, 2018 for a review). For example, in a sample of experienced golfers, Moore and colleagues (2013) found that cognitive evaluations more consistent with a challenge state were related to better performance than evaluations more indicative of a threat state (Moore et al., 2013). Thus, knowing how to promote a challenge state (or counteract a threat state) could enable the optimisation of athletic performance during pressurized competition. Related to this notion, the TCTSA specifies that high self-efficacy, high perceived control, and an approach focus promote more favourable cognitive evaluations and a challenge state. The TCTSA also specifies that a challenge state leads to more efficient attention, positive emotions, and emotions being perceived as more facilitative for performance (Jones et al., 2009). In contrast, low self-efficacy, low perceived control, and an avoidance focus promote less favourable cognitive evaluations and a threat state. According to the TCTSA, a threat state results in less efficient attention (i.e., a focus on task-irrelevant stimuli), negative emotions, and emotions being perceived as unhelpful for performance (Jones et al., 2009).

Previous laboratory-based research has successfully manipulated CAT states either directly with scripts influencing evaluations of situational demands and/or personal coping resources (e.g., verbal instructions, Moore, Vine, Wilson, & Freeman, 2012; audio instructions, Turner, Jones, Sheffield, & Barker, 2014), or indirectly via psychological interventions (e.g., arousal reappraisal, Moore, Vine, Wilson, & Freeman, 2015; quiet eye training, Moore, Vine, Freeman, & Wilson, 2013; imagery, Williams & Cumming, 2012). Despite some promising findings demonstrating the successful manipulation of CAT states and performance (e.g., study 2, Feinberg & Aiello, 2010; Moore et al., 2013; Moore et al., 2015[8]), other evidence has been more equivocal. Indeed, in one study, the manipulation only had a marginally significant effect on CAT states, and the threat group outperformed the challenge group (i.e., study 1, Feinberg & Aiello, 2010). Meanwhile, in the two other studies, the manipulation check confirmed a successful manipulation of underlying demand and resource evaluations (study 4, Feinberg & Aiello, 2010; Williams & Cumming, 2012), but there were no effects on task performance. Following these mixed findings, it is important to examine if other psychological interventions can lead to a challenge state and improved performance. One possible intervention is self-talk.

[8] Moore et al. (2015) reported a difference on CAT states that did not reach statistical significance (p = .17), but can be considered practically significant as it equated to a medium effect size (d = 0.44). The performance difference between the experimental groups was statistically, as well as practically, significant (p = .02 , d = 0.93)

Self-talk is often used in sport to direct attention, create more positive interpretations of anxiety, and optimise performance (Hatzigeorgiadis, Zourbanos, Galanis, & Theodorakis, 2011; Wadey & Hanton, 2008). Self-talk includes spontaneously occurring automatic thoughts and verbalisations, and deliberate and strategic statements addressed to oneself (Hardy et al., 2009). Self-talk can vary in terms of content, emotional valence, and whether it is audible or silent and deliberate or automatic (Theodorakis, Weinberg, Natsis, Douma, & Kazakas, 2000; Theodorakis, Hatzigeorgiadis, & Zourbanos, 2012; van Raalte, Vincent, & Brewer, 2016). A recent review distinguished organic and strategic self-talk, which represent self-statements reflecting ongoing cognitive processes and cue words used for strategic purposes, respectively (Latinjak, Hatzigeorgiadis, Comoutos, & Hardy, 2019). Organic self-talk has further been divided into spontaneous and goal-directed self-talk, which represent the unintentional (automatic) and intentional responses to athletes' emotions and thoughts. Beyond these distinctions, two of the most common forms of self-talk are instructional (i.e., cues that direct attention and instruct regarding technical, strategic, or kinaesthetic aspects of skill execution) and motivational (i.e., cues that maximise motivation, effort, confidence, and positive mood; Hatzigeorgiadis et al., 2011). A systematic review found that both forms of self-talk improved performance (Tod, Hardy, & Oliver, 2011). Motivational self-talk has also been shown to reduce cognitive anxiety and enhance self-confidence (Hatzigeorgiadis, Zourbanos, Mpoumaki, & Theodorakis, 2009).

Furthermore, a key self-talk theoretical model, the framework for the study and application of self-talk within sport (Hardy et al., 2009), specifies that self-talk can exert effects on attention, motivation, affect, and behaviour in ways similar to a challenge state. Specifically, self-talk is thought to improve concentration and reduce interfering thoughts, increase self-efficacy, improve anxiety and anxiety interpretations, and optimize movement and skill execution. However, none of the abovementioned theories specify CAT states as a potential mechanism in the relationship between self-talk and performance. Thus, a study examining the effect of self-talk on CAT states could significantly contribute to the literature.

As theoretical models and empirical research in the CAT and the self-talk literature propose consistent effects of a challenge state and effective self-talk (i.e., improved performance, attention, self-efficacy, and more facilitative interpretations of emotions), the present study aimed to examine whether self-talk directly influenced CAT states. We hypothesised that in anticipation of a post-training dart-throwing task, participants in the instructional and motivational self-talk groups would report cognitive evaluations (i.e., coping resources match/exceed task demands), and exhibit cardiovascular responses (i.e., relatively higher CO and/or lower TPR reactivity), more reflective of a challenge state than those in a control self-talk group (verbalising the trial number as a neutral self-talk cue; H1). Furthermore, we hypothesised that participants in the instructional and motivational self-talk groups would perform a post-training dart-throwing task better than those in a control self-talk group (relative to pre-training performance; H2). Finally, we hypothesised that cognitive evaluations (i.e., coping resources match/exceed task demands), and cardiovascular responses (i.e., relatively higher CO and/or lower TPR reactivity), more consistent with a challenge (versus a threat) state would be related to better task performance (H3). The hypothesised relationships between self-talk, CAT states, and performance are graphically illustrated in Figure 5.1.

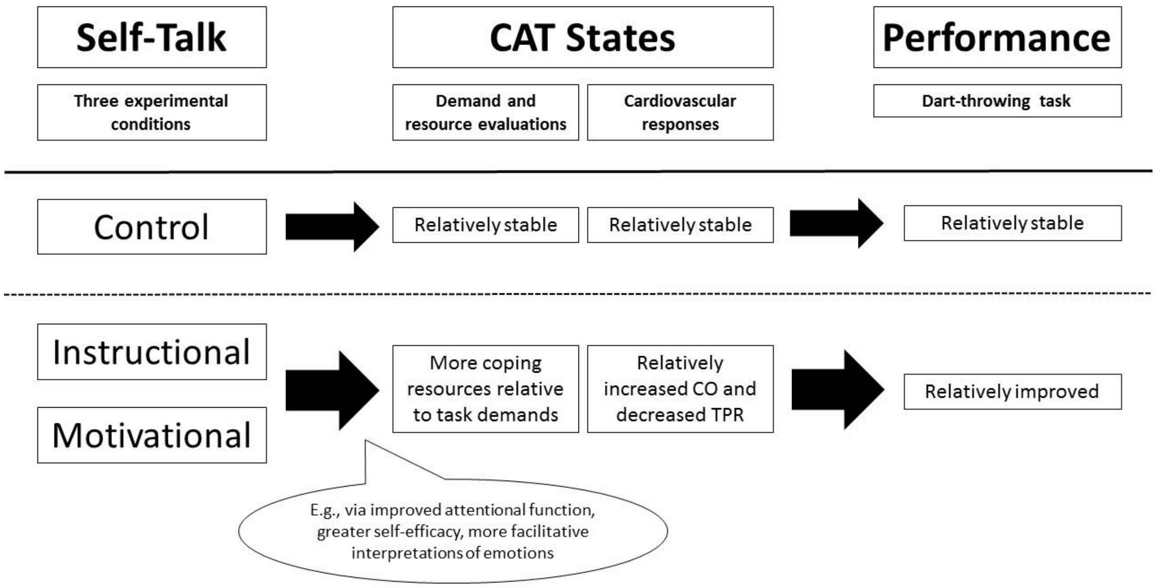

Figure 5.1. Hypothesised relationships between self-talk, CAT states, and performance.

5.3 Method

5.3.1 Participants

A power calculation for a repeated-measures ANOVA with a between-within interaction was conducted using G*Power software version 3.1.9.2. Because no effect size could be obtained for the effect of self-talk on CAT states, a medium effect size was assumed (d = 0.50; Cohen, 1992). This is consistent with the average effect of self-talk on performance (d = 0.48; Hatzigeorgiadis et al., 2011). With an alpha level of 0.05, and 90% desired power, the power calculation produced a minimum sample size of 54 (60 for d = 0.48). The final sample consisted of 62 university students and members of staff (84% male; M_{age} = 24 years, SD = 6, range 18-52). Native English speakers comprised 55% of the sample[9]. All participants reported being right-handed or ambidextrous. Two

[9] The main analyses were repeated to control for potential effects of native language (coded dichotomously for English versus non-English). This showed no significant effects for native language, and did not

participants reported having played darts at club level, whereas the remaining participants reported not engaging in competitive darts before.

5.3.2 Materials

5.3.2.1 Cardiovascular data. The Portapres Model-2 (Finapres Medical Systems BV, Amsterdam, the Netherlands) was used to record three cardiovascular variables: HR, CO, and TPR. The Portapres bases its measurements on the arterial volume-clamp method of Peñáz (1973), and the physiological calibration criteria for the proper unloading of the finger arteries of Wesseling (1996). It also uses a height correction unit to compensate for hydrostatic pressure changes due to movement of the hand. Previous research has used the Portapres for CAT measurements (e.g., Hase, Gorrie-Stone, & Freeman, 2018; Moore, Young, Freeman, & Sarkar, 2018), and it has been validated against the Finapres and Oxford method, and was found to be accurate, reliable, and cause no more missing data due to artefacts than the latter method (Hirschl, Woisetschläger, Waldenhofer, Herkner, & Bur, 1999; Imholz et al., 1993). Data were converted and downloaded for analysis using Beatscope software version 1.1.

5.3.2.2 Demand and resource evaluations. Demand and resource evaluations were assessed via two self-report items from the Stressor Appraisal Scale (Schneider, 2008). These items have been well-established in the CAT literature, and have been used to validate CAT cardiovascular indices (e.g., Tomaka, Blascovich, Kibler, & Ernst, 1997; Tomaka et al., 1993), and in research linking cognitive evaluations, cardiovascular responses, and performance (e.g., Hase, Gorrie-Stone, et al., 2019; Vine et al., 2013). Specifically, these items asked participants: "How demanding do you expect the

---

change the general pattern of results.

upcoming task to be?” and “How able are you to cope with the demands of the upcoming task?”. Consistent with Schneider (2008), both items were scored on a seven-point Likert scale anchored between *not at all* (1) and *extremely* (7). A cognitive CAT variable (i.e., demand resource evaluation score; DRES) was then created by subtracting evaluated demands from resources, meaning that scores ranged from -6 to 6 and higher values denoted evaluations more consistent with a challenge state (i.e., resources match/exceed demands; Moore et al., 2013).

5.3.2.3 Self-talk manipulation check. Two self-report items were used to ask participants about their self-talk use: “How often did you repeat your self-talk statement?” and “Do you believe that this procedure was helpful to you?” (Theodorakis et al., 2000). Both items were scored on a 10-point scale anchored between *not at all* (1) and *extremely* (10).

5.3.2.4 Dart-throwing performance. Participants threw darts from a distance of 2.4 m toward a dartboard of 44.8cm diameter, with the centre (bulls-eye) 1.7m above the floor. Unlike a traditional dartboard, the board was divided into nine concentric circles around a red bulls-eye. Landing a dart in the outermost ring was worth one point, with every more central ring worth one more point, and 10 points being awarded for landing the dart in the bulls-eye. Darts that landed outside the outermost ring scored zero points. Time to complete each task was recorded, although there was no time limit for this task[10].

[10] Time required to complete the task was not significantly different between groups for the baseline task [$F(2, 59) = 0.36$, $p = .70$, $\eta_p^2 = .01$], nor the final task [$F(2, 59) = 0.44$, $p = .65$, $\eta_p^2 = .02$].

5.3.3 Procedure

This study was approved by an institutional ethics committee. Upon entering the laboratory, participants were given an information sheet and provided informed consent. The information sheet explained the study and highlighted that rewards would be given to the three best performers on the two pressurised dart-throwing tasks (i.e., combined baseline and final score), which each task consisting of 20 throws. The study protocol including the order of the dart-throwing tasks is illustrated in Figure 5.2 and comprised: (1) baseline task (20 throws), (2) first training block (10 throws), (3) second training block (10 throws), and (4) final task (20 throws). Before starting the baseline task, participants sat in front of a computer screen and a Qualtrics survey guided them through the study protocol. Participants first provided demographic information (e.g., age, sex, native language, previous darts experience), and then the experimenter put the Portapres on the left hand of participants (cardiovascular measurements with this device may be sensitive to laterality, which is why right-handed or ambidextrous participants were recruited), with the cuff around the middle finger and the height correction sensor around the upper arm at the height of the sternum. Resting cardiovascular data were then recorded for three minutes (as Vine, Freeman, Moore, Chandra-Ramanan, & Wilson, 2013), after which task instructions were shown on the computer screen. Participants were asked to confirm that they had read the instructions, and then think about the instructions and the upcoming task for one minute, during which cardiovascular data was recorded. Participants then reported demand and resource evaluations before standing up and performing the baseline task (20 throws). Performance was recorded for all throws.

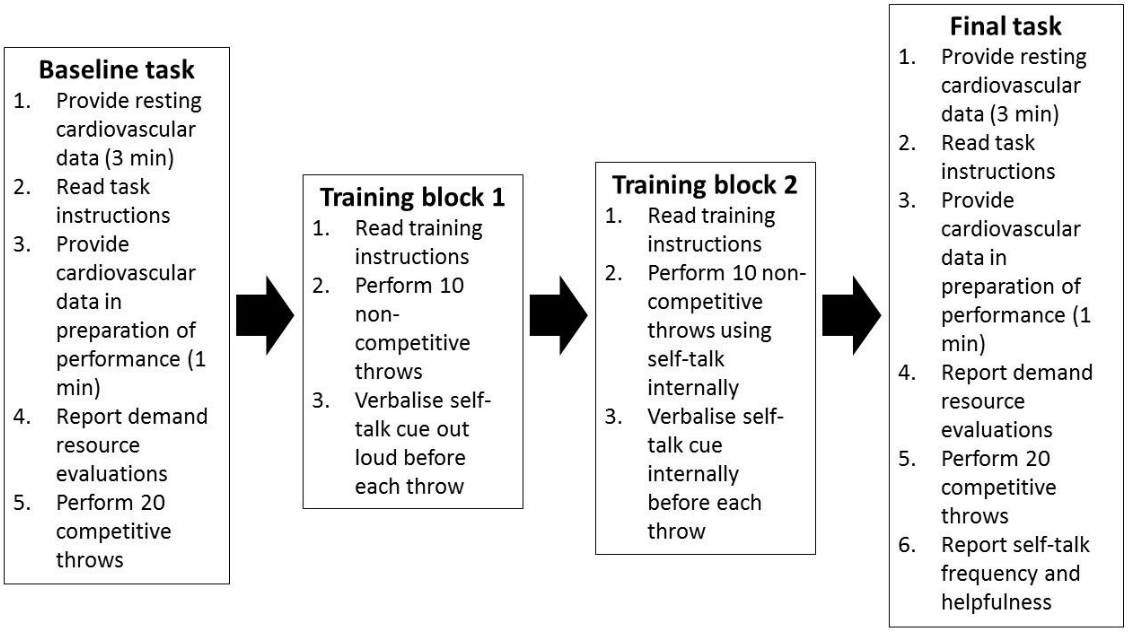

Figure 5.2. Study protocol and order of dart-throwing tasks.

Next, participants were randomly assigned (with a randomiser embedded in the Qualtrics survey) to the instructional, motivational, or control self-talk group, and received instructions on the screen to stand up and perform the first training block comprising 10 throws. Immediately before each of these throws, participants verbalised their self-talk cue out loud. The self-talk cues were adapted from Theodorakis et al. (2000), who used the same motivational self-talk cue (i.e., “I can”). Due to the different tasks used in their studies, we modified the instructional self-talk cue to maintain a visual attentional focus on the target of the dart-throwing task (i.e., “aim central”; aiming to promote a quiet eye; Moore et al., 2013). In the control self-talk group, the self-talk cue was “Trial *x*”, where *x* stands for the number of the throw. It was emphasised that these throws were for training purposes only, and that the scores would not contribute to the final competitive score. After the first training block, participants were instructed to perform another 10 training throws in a second block, this time verbalising the self-talk cue internally before each throw. Once participants had completed the second training block, they were seated in front of the computer screen again and underwent another cardiovascular measurement with the same procedure as the first one (i.e., three minutes of rest, receipt of task instructions, and one minute reflection after task instructions). The instructions reminded them that their final task performance would count toward their final competitive score. After the cardiovascular recording had ended, participants reported demand and resource evaluations, stood up, and completed the final dart-throwing task (20 throws). Participants then sat down in front of the computer screen to complete the self-talk manipulation check items before they were debriefed and thanked.

5.3.4 Statistical Analysis

Mean heart rate (HR), cardiac output (CO), and total peripheral resistance (TPR) values were calculated for the final minute of the rest period and the one minute after task instructions for both the baseline and final dart-throwing tasks. Six univariate outliers[11] (values more extreme than three standard deviations from the mean) were winsorised to be 1% more extreme than the next non-outlying score (as Hase, Gorrie-Stone, et al., 2018). Resting CO and TPR values were then regressed on their respective post-instruction values with the standardised residuals saved to create residualised change scores that adjusted for baseline differences (Burt & Obradović, 2013). TPR residualised change scores were then multiplied by -1 and summed with the CO residualised change scores to create a single cardiovascular challenge and threat index (i.e., CTI), with a higher CTI representing a cardiovascular response more indicative of a challenge state (i.e., relatively higher CO and/or lower TPR reactivity).

[11] For each task, two outliers were winsorised on rest TPR, and one on post-instruction TPR.

As is common in CAT research (e.g., Vine et al., 2013), paired-samples t-tests were used to examine whether the sample as a whole were engaged in the task, by comparing resting and post-instruction HR on the baseline and final task, respectively. To check self-talk compliance and perceived helpfulness between the groups, two one-way between-subjects ANOVAs compared differences between the self-talk groups in terms of self-talk frequency and helpfulness. Simple contrasts with the control group as the reference group probed significant effects for self-talk group.

To test H1, two repeated-measures ANOVAs examined demand resource evaluation score (DRES) and CTI with task (i.e., baseline versus final) as the within-participants factor, and the group by task interaction as the between-participants factor and independent variable of interest. To explore significant effects, simple contrasts were used with the control self-talk group as the reference group.

H2 and H3 were tested with Generalised Estimating Equations (GEE) analysis predicting performance with self-talk group, task (i.e., baseline versus final), DRES, CTI, and the respective two-way interaction terms for task and self-talk group[12]. Specifically, H2 was tested with the group by task interaction effect, comparing the self-talk groups on change in performance from the baseline to the final task. Moreover, H3 was tested with the main effects for DRES and CTI on performance across tasks and groups. The GEE model was used because it enables a test of the relationships between a set of categorical and continuous independent variables (including their interactions), and a dependent variable across different time points, which is a parsimonious alternative to conducting

[12] i.e., group by task, group by cognitive CAT, group by cardiovascular CAT, task by cognitive CAT, and task by cardiovascular CAT.

separate analyses at each time point. All of the above analyses used a significance level of $\alpha = .05$.

5.4 Results

5.4.1 Preliminary Analyses

One participant provided no DRES for the final task, and the equipment did not record cardiovascular data for 10 participants due to signal problems[13]. The paired-samples t-tests for HR showed increases for both competitive tasks, although the difference was only marginally significant for the baseline task [M_{Baseline} = 1.38 bpm, 95% CI (-0.04; 2.79), $t(53) = 1.95$, $p = 0.06$, $d = 0.27$; M_{Final} = 2.24 bpm, 95% CI (0.32; 4.16), $t(52) = 2.34$, $p = 0.02$, $d = 0.32$]. Tables 5.1 and 5.2 provide descriptive statistics for DRES, CTI, performance, self-talk frequency, and self-talk helpfulness by self-talk group and task. The ANOVA on self-talk frequency revealed no significant difference between the groups [$F(2, 55) = 0.78$, $p = 0.46$, $\eta_p^2 = .03$], with the descriptive statistics indicating that participants in all groups almost always used their respective self-talk cues (see Table 5.1). The ANOVA on the self-talk helpfulness variable revealed a significant difference between the groups [$F(2, 55) = 3.43$, $p = 0.04$, $\eta_p^2 = .11$]. Simple contrasts indicated that the motivational group rated their self-talk cue to be significantly more helpful than the control group (contrast value = 1.75, $p = 0.01$), whereas the instructional group rated their self-talk cue to be more helpful than the control group, albeit not significantly so (contrast value = 1.21, $p = 0.09$)[14].

[13] One participant missed baseline task data, two participants missed final task data, and seven participants missed data from both tasks. Hence, the final sample comprised 61 participants for analyses of DRES and 52 participants for analyses of CTI.
[14] Changing the reference group revealed that the motivational and instructional self-talk groups were not significantly different on self-talk frequency, nor on self-talk helpfulness.

Table 5.1

Variables of Interest by Self-Talk Group and Task

	Instructional Self-Talk				Motivational Self-Talk				Control Self-Talk			
	Baseline Task		Final Task		Baseline Task		Final Task		Baseline Task		Final Task	
	M	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
1. Performance	114.25	16.35	121.95	14.98	118.45	21.41	127.68	22.14	127.10	17.35	129.70	13.93
2. DRES	1.90	2.00	2.40	2.25	2.66	1.74	2.89	2.14	2.53	1.85	2.85	1.66
3. CTI	0.18	2.04	-0.25	1.02	0.27	1.50	-0.14	2.02	-0.55	1.73	0.44	1.88
4. Self-Talk Frequency	N/A	N/A	7.58	2.59	N/A	N/A	8.55	1.96	N/A	N/A	8.16	2.71
5. Self-Talk Helpfulness	N/A	N/A	6.16	1.83	N/A	N/A	6.70	2.11	N/A	N/A	4.95	2.41

Note. DRES = Demand resource evaluation score. CTI = Challenge and threat index.

Table 5.2

Raw Cardiovascular Variables by Self-Talk Group and Task

	Instructional Self-Talk				Motivational Self-Talk				Control Self-Talk			
	Rest		Post-instructions		Rest		Post-instructions		Rest		Post-instructions	
Baseline Task	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
1. HR (bpm)	77.49	13.30	80.87	13.98	81.91	14.72	82.30	14.97	78.76	10.15	79.30	9.65
2. CO (lpm)	5.44	1.96	5.78	1.81	6.03	2.46	6.46	2.31	5.83	1.40	5.90	1.80
3. TPR (mmHg.s/ml)	1.02	0.37	0.92	0.23	0.92	0.49	0.86	0.37	0.94	0.36	0.93	0.32
Final Task	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
4. HR (bpm)	77.54	12.84	81.35	13.50	81.31	12.67	82.79	14.59	77.48	9.31	79.14	11.91
5. CO (lpm)	5.83	1.73	5.89	1.46	6.09	2.20	6.13	2.29	5.43	1.40	5.98	1.71
6. TPR (mmHg.s/ml)	0.96	0.38	1.01	0.50	0.95	0.49	0.98	0.61	0.91	0.20	0.91	0.19

Note. HR = Heart rate. CO = Cardiac output. TPR = Total peripheral resistance.

5.4.2 Main Analyses

5.4.2.1 H1: Effects of self-talk manipulations on CAT states. Table 5.3 summarises the two repeated-measures ANOVAs on DRES and CTI. There were no significant effects for self-talk group by task on DRES [$F(2, 58) = 0.97$, $p = .39$, $\eta_p^2 = .03$], or CTI [$F(2, 49) = 1.59$, $p = 0.21$, $\eta_p^2 = .06$]. Despite the lack of statistical significance, these baseline-to-final task changes represented small and medium effect sizes, respectively.

Table 5.3

Mixed-Model ANOVAs on DRES and CTI by Self-Talk Group

	DRES				CTI			
	Mean Square	*F*	*p*	η_p^2	Mean Square	*F*	*p*	η_p^2
Task	2.02	3.31	.07	.05	0.00	0.00	< .99	.00
Self-Talk Group	0.59	0.97	.39	.03	5.52	1.59	.21	.06
Error	0.61				3.46			

Note. DRES = Demand resource evaluation score. CTI = Challenge and threat index.

5.4.2.2 H2: Effects of self-talk manipulations on performance. Table 5.4 presents parameter estimates for the GEE analysis predicting performance relevant to H2 and H3. There was a significant group by task interaction effect (Wald $\chi^2 = 6.11$, $p = .05$). The parameter estimates for this effect showed that the performance of the motivational group improved more from the baseline to the final task than the performance of the control group ($B = -11.76$, Wald $\chi^2 = 5.52$, $p = .02$), but there was no significant difference in performance change from the baseline to the final task between the instructional and control groups ($B = -3.36$, Wald $\chi^2 = 0.38$, $p = .54$).

Table 5.4

GEE Analysis Parameter Estimates

Effect	Comparison	*B*	Wald χ^2	*p*
Main Effects				
Self-Talk Group				
	IST – CST	-9.62	2.70	.10
	MST – CST	-7.94	1.14	.29
Task				
	BL – FT	-0.21	0.00	.96
DRES	N/A	2.64	4.37	.04
CTI	N/A	-0.31	0.18	.67
Interaction Effects				
Self-Talk Group by Task				
	$(IST_{BL} – CST_{BL}) – (IST_{FT} – CST_{FT})$	-3.36	0.38	.54
	$(MST_{BL} – CST_{BL}) – (MST_{FT} – CST_{FT})$	-11.76	5.52	.02
DRES by Self-Talk Group				
	$DRES_{IST}$ - $DRES_{CST}$	-1.89	1.17	.28
	$DRES_{MST}$ - $DRES_{CST}$	1.37	0.63	.43
CTI by Self-Talk Group				
	CTI_{IST} - CTI_{CST}	-4.62	6.35	.01
	CTI_{MST} - CTI_{CST}	2.01	3.74	.05
DRES by Task				
	$DRES_{BL}$ - $DRES_{FT}$	0.37	0.18	.68
CTI by Task				
	CTI_{BL} - CTI_{FT}	2.61	4.84	.03
Intercept		126.59	605.86	.00

Note. BL = Baseline task. FT = Final task. CST = Control self-talk. IST = Instructional self-talk. MST = Motivational self-talk. DRES = Demand and resource evaluation score. CTI = Challenge and threat index. N/A = No applicable comparison due to the continuous nature of the variable.

5.4.2.3 H3: Effects of CAT states on performance. There was a significant main effect for DRES (Wald $\chi^2 = 13.33$, $p < .01$). Furthermore, there were significant interaction effects for group by CTI (Wald $\chi^2 = 11.54$, $p < .01$), and for task by CTI (Wald $\chi^2 = 4.84$, $p = .03$). Parameter estimates for the DRES main effect showed that DRES more consistent with a challenge state (i.e., coping resources match/exceed task demands) was associated with better performance ($B = 2.64$, Wald $\chi^2 = 4.37$, $p = .04$). The parameter estimates for the group by CTI interaction effect showed group differences in the way CTI related to performance. Specifically, CTI was significantly more negatively related to performance for the instructional group than the control group ($B = -4.62$, Wald $\chi^2 = 6.35$, $p = .01$). In contrast, CTI was marginally more positively related to performance for the motivational group than the control group ($B = 2.01$, Wald $\chi^2 = 3.74$, $p = .05$). Hence, CTI more consistent with a challenge state (i.e., relatively higher CO and/or lower TPR reactivity) was more favourable for the motivational group than the control group, and in turn for the control group than the instructional group. Finally, the parameter estimate for the task by CTI interaction effect showed that CTI was more positively related to performance in the baseline task than in the final task ($B = 2.61$, Wald $\chi^2 = 4.84$, $p = .03$).

5.5 Discussion

This study examined the effects of self-talk on CAT states and performance during a competitive dart-throwing task. It was predicted that (H1) the instructional and motivational self-talk groups would exhibit cognitive evaluations and cardiovascular responses more indicative of a challenge state compared to the control group, (H2) the instructional and motivational self-talk groups would perform the final task better (relative to baseline) than the control group, and (H3) both cognitive evaluations and cardiovascular responses more indicative of a challenge state would be related to better performance (see Figure 5.1). H1 was not supported, but there was partial support for H2, as participants in the motivational self-talk group improved their performance from the baseline to the final task more than participants in the control group. There was also partial support for H3, as demand and resource evaluations more consistent with a challenge state were related to better performance. Hence, this study provides initial insight into the relationships between self-talk, CAT states, and task performance. Figure 5.3 details the actual findings of this study in contrast to the findings hypothesised in Figure 5.1.

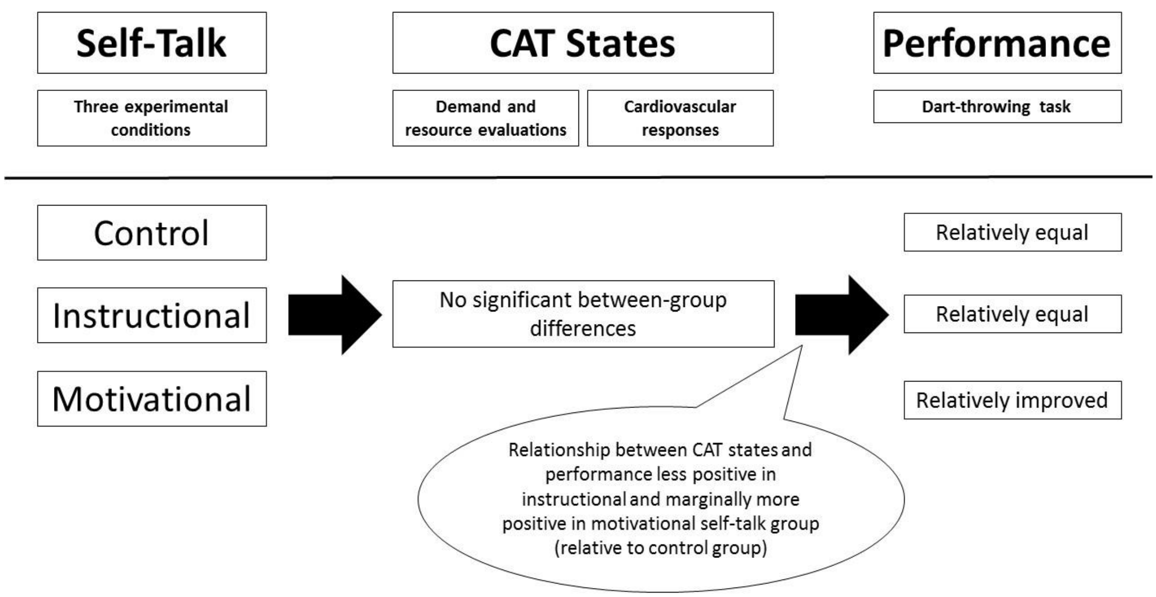

Figure 5.3. Overview of main results

Instructional and motivational self-talk, as practiced in this study, did not significantly affect CAT states, assessed at both the cognitive and cardiovascular level. Indeed, the differences in how the groups changed from baseline to final task represented small (DRES) and medium (CTI) effects, which was smaller than (DRES) and about equal to (CTI) the effect size assumed in the power calculation. As this study is the first to investigate this relationship, there is no previous evidence regarding the association between self-talk and CAT states. However, previous research and theory has linked instructional and motivational self-talk with constructs that have also been linked with CAT states including performance, attentional focus, goal orientation, and interpretations of anxiety (e.g., Hardy et al., 2009; Hatzigeorgiadis et al., 2009; Hatzigeorgiadis et al., 2011; Jones et al., 2009; Latinjak, Torregrossa, Comoutos, Hernando-Gimeno, & Ramis, 2019; Vine, Moore, & Wilson, 2016). Given the current findings, it appears that effective self-talk does not directly influence CAT states, despite this apparent consistency. However, self-talk interventions might be more effective if they deliberately focus on the antecedents of CAT states proposed by the TCTSA (e.g., self-efficacy, perceived control, and/or achievement goals). For example, Williams and Cumming (2012) elicited different CAT appraisals of a dart-throwing task via imagery scripts that focused on self-efficacy, perceived control, and achievement goals. Furthermore, self-talk interventions might be more beneficial if they consider important moderators. For example, Hardy and colleagues (2009) proposed that belief in self-talk and cognitive-processing preference might moderate the effectiveness of self-talk. Thus, it could be that self-talk only promotes a challenge state for individuals who believe that self-talk is effective, or those with a verbal cognitive processing preference.

Motivational self-talk, as practiced in this study, was found to enhance dart-throwing performance. Specifically, the motivational self-talk group demonstrated greater improvements in performance from the baseline to the final task than the control group. This trend was also present for the instructional group, but it did not reach statistical significance. As such, these results are not fully consistent with the findings of systematic reviews and meta-analyses, which have found that both instructional and motivational self-talk benefit performance (Hatzigeorgiadis et al., 2011; Tod et al., 2011). A potential explanation for the differences between the experimental groups (relative to the control group) observed in the present study, could be the perceived helpfulness of the self-talk cue, as the motivational, but not the instructional group, rated their cue to be more helpful than the control group. As Hardy and colleagues (2009) mentioned that efficacy beliefs about self-talk can moderate the relationship between self-talk and task performance, an instructional self-talk cue appearing more helpful to participants might also have produced a significant improvement in performance relative to the control group.

It is worth noting that the control group in this study differed from some control groups in previous studies. For instance, some control groups have received no self-talk instructions at all (i.e., no-verbalisation controls; e.g., Hatzigeorgiadis et al., 2009). In contrast, this study used a control self-talk cue to impose similar cognitive load on participants and to prevent organic self-talk, which may occur in no-verbalisation controls (e.g., Hardy, Hall, Gibbs, & Greenslade, 2005). Although such a condition could theoretically function as a negative intervention (i.e., hampering adaptive organic self-talk use), it appears that this was not the case in this study, as the DRES and CTI data (Table 5.1) suggested that the control self-talk group exhibited a trend toward cognitive evaluations and cardiovascular responses more consistent with a challenge state than the instructional and motivational self-talk groups. As silence in a pressurised situation might also provoke organic self-talk focused on maladaptive cognitions and/or attentional processes, the control self-talk cue in this study (i.e., *Trial one,* etc.) might have had a protective effect, distracting participants from task-irrelevant stimuli, and refocusing them on the task before every throw. However, for participants who habitually use organic self-talk in an adaptive way, the control self-talk cue might have had a negative effect by disrupting organic self-talk.

In this study, cognitive evaluations more indicative of a challenge state (i.e., coping resources match/exceed task demands) were related to better performance. This is consistent with the predictions of the BPSM and TCTSA (Blascovich, 2008; Jones et al., 2009), and the findings of a recent systematic review, in which 76% of the reported effects found that a challenge evaluation was associated with better performance than a threat evaluation (Hase, O'Brien, et al., 2018). In contrast, CTI had no significant effect on task performance. This lack of association is inconsistent with the predictions of the BPSM and TCTSA, and the findings of recent reviews (e.g., Behnke & Kaczmarek, 2018), although some studies assessing both cognitive and cardiovascular measures of CAT states have also found divergent effects (e.g., Moore et al., 2018; Vine et al., 2013). Correlations between cognitive and cardiovascular measures of CAT states are usually weak to moderate (e.g., Moore et al., 2018; Vine et al., 2013), and the correlation between DRES and CTI in this study was not significant, raising concerns about the predictions of the BPSM.

Rather than self-talk influencing CAT states, they might operate in an interactive manner, as this study observed an interaction effect between CTI and self-talk on task performance. Specifically, CTI was less positively related to performance in the instructional than in the control self-talk group. One explanation for this would be instructional self-talk being more beneficial for individuals in a threat state than for individuals in a challenge state. Precisely, the instructional self-talk cue might have promoted a more optimal attentional focus on the target, which is absent in a threat state, but already present in a challenge state, as it is one of the proposed mechanisms with which a challenge state is thought to operate (see Vine, Moore, & Wilson, 2016). In this vein, the TCTSA proposes that "in a challenge state the focus of attention is on appropriate cues, whereas in a threat state attention is also directed to task irrelevant stimuli that could cause harm" (Jones et al., 2009, p. 173). Hence, the direction of attention towards the target in the instructional group may not have benefited those in a challenge state (who should have focused on the target anyway), but may have helped those in a threat state (who should have otherwise focused on task-irrelevant cues). As a result, CTI may have impacted performance less strongly in the instructional than in the control and motivational self-talk groups.

In addition to the result noted above, there was a more positive relationship between CTI and performance in the motivational than in the control self-talk group, although this effect only approached significance. While not significant, this potentially meaningful descriptive trend suggests that motivational self-talk might have offered the most benefit to those who responded to the task with a cardiovascular response more indicative of a challenge state (i.e., relatively higher CO and/or lower TPR reactivity), or that it had a counterproductive effect on participants who responded to the task with a pattern more reflective of a threat state. While speculative, a potential explanation for this result could be that motivational self-talk encouraged more liberal use of available energy (e.g., by increasing effort), which could have conflicted with the cardiovascular threat pattern, as energy mobilisation is not very efficient in a threat state (due to little cardiac activity and/or vasoconstriction). Conversely, it could work synergistically with the cardiovascular challenge pattern due to the more efficient energy mobilisation in a challenge state (due to greater cardiac activity and/or vasodilation, Blascovich, 2008).

Some limitations should be noted. First, the self-talk intervention was very brief and had a low self-determination component (Hardy, 2006). Ideally, the selection of self-talk cues should have been determined by assessing individual needs and preferences (e.g., whether to verbalise cues aloud or internally; Hatzigeorgiadis, Zourbanos, Latinjak, & Theodorakis, 2014), selecting individually matching cues, and adapting, internalising, and automatizing cues in training (Hardy, 2006). Also, the self-talk cues were only aimed at a subset of the functions covered by more complete interventions of the same type (e.g., "*I can*" targets confidence, but not effort or arousal control; "*Aim central*" directs attention, but does not introduce technical information or influence decision-making), and therefore may not have been sufficient to elicit changes in CAT states. Future research could therefore test a prolonged self-talk intervention covering multiple testing sessions.

Second, due to the lack of a no-verbalisations control group, it is difficult to infer whether the improvements in performance from the baseline to the final task were attributable to practice effects, an effect of all three self-talk cues, or both. Furthermore, the control self-talk cue might have impacted organic self-talk and thereby CAT states and performance. Although there might not have been a negative impact on CAT states (see Table 5.1), future research should include both a control self-talk and a no-verbalisations condition, and obtain reports of cognitive load and organic self-talk use to provide conclusive evidence to answer this question. Similarly, the manipulation check used in this study did not assess organic self-talk, which might have been assessed in parallel to the strategic self-talk that participants used (Latinjak, Hatzigeorgiadis, et al., 2019).

Third, in the baseline task, task engagement was relatively weak, as evidenced by the marginally significant increase in HR. Future research might be able to prevent this by verbally and emphatically delivering task instructions, and/or provoking elevated pressure by highlighting social comparison (e.g., being filmed, mentioning a scoreboard) or performance-contingent punishments (e.g., being interviewed for poor performance; Moore et al., 2015). Other studies that have observed greater increases in HR, however, have compared a quiet rest period to a more metabolically demanding period (e.g., a speech; Blascovich, Seery, Mugridge, Norris, & Weisbuch, 2004). Thus, the silent task visualisation in the present study might have produced cardiovascular data that were more reflective of purely psychological task engagement, rather than speech production or other factors reflecting physiological load.

5.5.1 Conclusion

This study examined the effect of self-talk on CAT states and performance during a competitive dart-throwing task. Self-talk did not impact CAT states, but motivational self-talk improved performance more than control self-talk. Thus, self-talk may be a useful psychological strategy, but it might not exert its beneficial effects on performance by influencing CAT states. In addition, a cognitive evaluation more reflective of a challenge state (i.e., coping resources match/exceed task demands) was related to better performance. Finally, the findings relating to the cardiovascular reactivity patterns of CAT states were more complicated, and suggested that instructional self-talk may weaken, whereas motivational self-talk may strengthen, the relationship between a challenge-like cardiovascular response (i.e., higher CO and/or lower TPR reactivity) and performance, compared to control self-talk. Hence, motivational self-talk may offer more benefit to athletes experiencing a challenge state, while instructional self-talk might be more advantageous to athletes in a threat state.

Table 5.5

Summary of Chapter 5 and Preview of Next Chapter

Chapter	Aim	Findings
5	To examine whether instructional and motivational self-talk promote a challenge state.	Neither instructional, nor motivational self-talk directly promoted a challenge state on the cognitive or cardiovascular CAT level. Instructional self-talk attenuated the relationship between CAT states and performance, whereas motivational self-talk exacerbated the relationship.
	Rationale for next chapter Since the self-talk interventions in chapter 5 did not promote a challenge state, another intervention study was conducted, this time on a previously untested physiological intervention. A tyrosine supplement was chosen as the focal intervention, as the BPSM specifies that catecholamine function is central to the occurrence of a challenge state, and tyrosine is a catecholamine precursor that has been shown to augment catecholamine function when ingested as a supplement.	
Chapter	Aim	Findings
6	To examine whether tyrosine intake promotes a challenge state.	

Chapter 6

Tyrosine Intake and Cardiovascular Responses in a Motivated Performance Situation

Published as:

Hase, A., Gorrie-Stone, T., & Freeman, P. (2019). Tyrosine intake and cardiovascular responses in a motivated performance situation. *Sport, Exercise, and Performance Psychology*, *8*, 80–92. doi: 10.1037/spy0000144.supp

6.1 Abstract

Ingesting the catecholamine precursor tyrosine can prevent decrements in, or improve, cognitive and motor performance in demanding situations. Furthermore, the biopsychosocial model of challenge and threat specifies that adrenal medullary catecholamine release plays a central role in the occurrence of a challenge state, which has been linked to better performance under pressure than a threat state. The present study thus examined whether acute tyrosine intake impacts upon challenge and threat states or influences cognitive and motor performance independently. A double-blind randomised crossover design with 49 participants (33 males; μ_{age} = 22.5 years, *SD* = 5.0) was used. Participants ingested tyrosine or placebo (150mg/kg body mass) 60 minutes before performing the N-Back task and a bean-bag throwing task. Cognitive self-reports and cardiovascular data before each task provided indicators of challenge and threat states. There were no significant differences between tyrosine and placebo on the cognitive and cardiovascular challenge and threat variables. GEE analyses found that tyrosine was associated with better performance than placebo on the bean-bag throwing task, but not on the N-Back task. A significant interaction effect showed that challenge and threat states were more positively related to performance in the placebo condition than in the tyrosine condition. This suggests that tyrosine may have attenuated the detrimental effect of a threat state. The present study breaks new ground in relating the impact of a dietary supplement to challenge and threat states and finding that tyrosine may in some cases attenuate the negative effects of a threat state.

6.2 Introduction

The question of why some individuals excel in important situations whereas others struggle under pressure is of great importance, and due to the widespread occurrence of situations in which active performance is required to attain a self-relevant goal, this topic is of interest to sport, social, organisational, and clinical psychologists alike. The BPSM of CAT (Blascovich, 2008a) is a key framework for understanding performance variation under pressure across these disciplines. It was extended and applied to the domain of sports by the TCTSA (Jones et al., 2009). In many studies, a challenge state has been associated with better performance than a threat state (for a review see chapter 2). This relationship has led researchers to study putative challenge-promoting interventions such as imagery, stress optimisation, and quiet eye training, and their effects on performance (Jamieson, Crum, Goyer, Marotta, & Akinola, 2018; Moore, Vine, et al., 2013; Williams & Cumming, 2012). These interventions typically aim at improving performance by optimising psychological antecedents of CAT states (e.g., self-efficacy, perceived control; Williams & Cumming, 2012), and by helping individuals interpret physiological arousal as more facilitative for performance (Jamieson et al., 2018; Moore, Wilson, et al., 2013). However, these interventions have all taken psychological approaches to manipulating CAT states. The current study therefore examined whether a nutritional intervention that targets a neurotransmitter group specified by the BPSM to be key to the occurrence of CAT states may promote a challenge state and enhance performance. Although some nutrients and supplements (e.g., sugar and caffeine; Grasser et al., 2016; Hartley, Lovallo, & Whitsett, 2004) exhibited effects on the cardiovascular system akin to those of CAT states, research examining dietary interventions in a CAT context is scarce.

The BPSM describes CAT states as responses that only occur in motivated performance situations, which are goal-relevant, evaluative, potentially stressful, and require sufficient active performance in order for personal growth (Blascovich & Mendes, 2000). CAT states differ in their underlying cognitive evaluations and concomitant physiological responses. A challenge state occurs when perceived personal coping resources outweigh or equal perceived situational demands, whereas a threat state occurs when perceived situational demands outweigh perceived personal coping resources. These demand and resource evaluations are thought to be influenced by several factors, such as self-efficacy, achievement goal orientation, perceived control, danger, uncertainty, novelty, required effort, skills, knowledge, abilities, presence of others, attitudes, and beliefs (Jones et al., 2009; Blascovich, 2008a). Physiologically, a challenge state has been hypothesised to involve an increase in sympathetic-adrenomedullary axis function. The sympathetic activation at the myocardium is thought to increase HR (the number of heart beats per minute) and stroke volume (the volume of blood ejected by the left ventricle with each heart beat) by acting on β1 receptors at the myocardium, thereby increasing CO (volume of blood ejected by the left ventricle per minute). At the same time, adrenal medullary release of epinephrine is thought to act as a vasodilator by acting on β2 receptors in skeletal muscle beds and bronchi, thereby decreasing TPR (the degree of systemic peripheral vascular constriction; Blascovich, 2008a; Blascovich & Mendes, 2000; Brownley et al., 2000).

In addition to sympathetic-adrenomedullary activation, a threat state is also thought to involve pituitary-adrenocortical axis activation that inhibits the sympathetic-adrenomedullary axis (Blascovich & Mendes, 2000). This leads to relatively small increases in HR, little change or minor decreases in CO, and little change or small increases in TPR during a threat state. The BPSM conceptualises CAT states as opposite ends to a bipolar continuum, meaning that one can be more or less strongly challenged or threatened, but not challenged and threatened at the same time. It also specifies task engagement, which is conceptualised as an increase in HR or VC (the contractile state of the left ventricle; operationalised by the BPSM as the inverse of the pre-ejection period), as a prerequisite for CAT states to occur in motivated performance situations. Hence, without task engagement neither a challenge nor a threat state will be experienced (Blascovich, 2008a).

Significant relationships between CAT states and performance have been found across diverse contexts. A recent systematic review of 38 studies that conceptualised CAT in a manner consistent with the BPSM found that a challenge state was related to better performance than a threat state in 28 of those studies (chapter 2). This relationship was generally supported regardless of CAT variable (cognitive, physiological, and dichotomous), outcome task (cognitive and behavioural), and research design used (correlational, quasi-experimental, direct experimental, and indirect experimental studies). For example, Turner and colleagues (2012) found that a physiological challenge state was related to better cognitive and motor task performance than a threat state, using a modified Stroop and a netball shooting task. Interestingly though, the available experimental studies only used psychological manipulations to induce CAT states. For example, some studies manipulated CAT with instructional sets targeting resource and demand evaluations (e.g., Feinberg & Aiello, 2010; Turner et al., 2014), and others targeted proposed psychological antecedents of CAT states (e.g., perceived required effort; Moore et al., 2014). The lack of physiological manipulations might be due to pioneering studies that successfully changed cardiovascular reactivity via manipulations of cognitive CAT evaluations, but did not succeed in evoking cognitive CAT evaluations via physiological manipulations, namely cold water immersion and physical exercise (Tomaka et al., 1997). To our knowledge, however, no study has examined the effects of a catecholamine-based intervention on CAT states. The BPSM of CAT specifies the catecholamine epinephrine to be centrally involved in the occurrence of a challenge state via stimulation of the vascular and cardiac epinephrine system (Blascovich & Mendes, 2000). Hence, a catecholamine-based CAT intervention could hold the potential to promote a challenge state and complement previous interventions. A possible catecholamine-based CAT intervention is supplemental tyrosine intake. The rationale for selecting tyrosine as an intervention in this study was that it is a catecholamine precursor whose consumption can affect catecholamine levels, and it has also exerted protective or enhancing effects on cognitive and motor performance under demanding conditions (Hase, Jung, & aan het Rot, 2015).

Tyrosine is a naturally occurring, non-essential amino acid. It is synthesised from phenylalanine and is converted into the dopamine precursor L-3,4-dihydroxyphenylalanine (L-DOPA) by the rate-limiting enzyme tyrosine hydroxylase. Tyrosine, but not its precursor phenylalanine, is able to stimulate catecholamine production in the brain, which has been observed directly and indirectly (for a review, see Fernstrom & Fernstrom, 2007). As tyrosine hydroxylase is usually about 75% saturated (Carlsson & Lindqvist, 1978), there is a modest, but significant potential to increase L-DOPA synthesis by increasing serum tyrosine levels, which should increase when demand is heightened due to greater neuronal activity (Fernstrom & Fernstrom, 2007). In the catecholamine pathway, tyrosine can be converted into L-DOPA, dopamine, and eventually norepinephrine and epinephrine. Importantly, an increase in serum tyrosine can be achieved through dietary supplementation. For example, Strüder et al. (1998) found that an acute dose of 10g of tyrosine significantly increased serum tyrosine levels in trained male cyclists within 45 minutes of ingestion. Importantly, they also found that the elimination half-life of tyrosine was sufficiently long for tyrosine levels to remain significantly elevated for 60 minutes following 150 minutes of cycling. Similarly, van de Rest and colleagues found that 150mg/kg body mass tyrosine ingestion led to a significant elevation in plasma tyrosine levels after 90 minutes, which persisted for another 150 minutes without substantial change to tyrosine levels (van de Rest, Bloemendaal, de Heus, & Aarts, 2017). Tumilty and colleagues found that 150mg/kg body mass of tyrosine significantly increased serum tyrosine levels within 60 minutes (Tumilty, Davison, Beckmann, & Thatcher, 2014). It should be noted, however, that other amino acids compete with tyrosine for uptake into the brain, and therefore it is advisable to administer tyrosine in a pure form and to restrict protein intake before administration in order to maximise brain tyrosine uptake (Fernstrom & Fernstrom, 2007).

The main mechanism of action by which tyrosine is thought to be effective is its stabilising influence on catecholamine levels in situations of heightened cognitive or physiological demands (e.g., cognitive load, extreme temperature), thereby preventing a performance decline. The importance of catecholamine function for cognitions, emotions, and behaviour has been demonstrated by depletion studies in which tyrosine and phenylalanine were removed from participants’ diet to elicit a depletion of brain catecholamine levels. Such a catecholamine depletion led individuals to behave in a less motivated manner (Cawley et al., 2013; McLean, Rubinsztein, Robbins, & Sahakian, 2004; Roiser et al., 2005), experience cognitive impairments (Harmer, McTavish, Clark, Goodwin, & Cowen, 2001), and become more susceptible to the detrimental effects of low light exposure (Cawley et al., 2013). Further, O'Brien and colleagues argued that catecholamine depletion may explain performance decrements in demanding situations, but that this may be mitigated by tyrosine consumption (O'Brien, Mahoney, Tharion, Sils, & Castellani, 2007). Indeed, a recent systematic review found that tyrosine intake protected or improved cognitive and motor performance under demanding conditions, while no beneficial effect was found for endurance exercise performance (Hase et al., 2015). For example, beneficial effects of tyrosine intake were found on reaction times following heat exposure (Kishore et al., 2013), on working memory performance following cold exposure (Mahoney, Castellani, Kramer, Young, & Lieberman, 2007; Shurtleff, Thomas, Schrot, Kowalski, & Harford, 1994), and on working memory performance under cognitive load (Thomas, Lockwood, Singh, & Deuster, 1999).

Given the previously presented work showing that 1) catecholamines are involved in CAT states (Blascovich, 2008a), 2) a challenge state generally relates to better performance than a threat state (chapter 2), 3) tyrosine intake can increase serum tyrosine and catecholamine levels (Fernstrom & Fernstrom, 2007), and 4) research has found tyrosine intake to improve cognitive and motor performance, we concluded that this evidence merits an examination of the impact of tyrosine on CAT states. Thus, the aim of the present study was to examine whether the beneficial effect of tyrosine intake on cognitive and motor performance is associated with a facilitation of a challenge state at physiological and psychological levels. We hypothesised that participants would exhibit relatively greater challenge reactivity (greater CAT index calculated from CO and TPR reactivity from baseline to post-task instructions) after tyrosine ingestion than after ingestion of a placebo (H1). In an exploratory manner, we also examined a potential effect of tyrosine on cognitive CAT evaluations. We also hypothesised that participants would perform better on a cognitive and a motor task after tyrosine ingestion than after placebo ingestion (H2). Finally, we hypothesised that a challenge state (measured as cardiovascular responses and cognitive evaluations) would be related to better performance than a threat state (H3). The hypothesised relationships between tyrosine intake, CAT states, and performance are graphically illustrated in Figure 6.1.

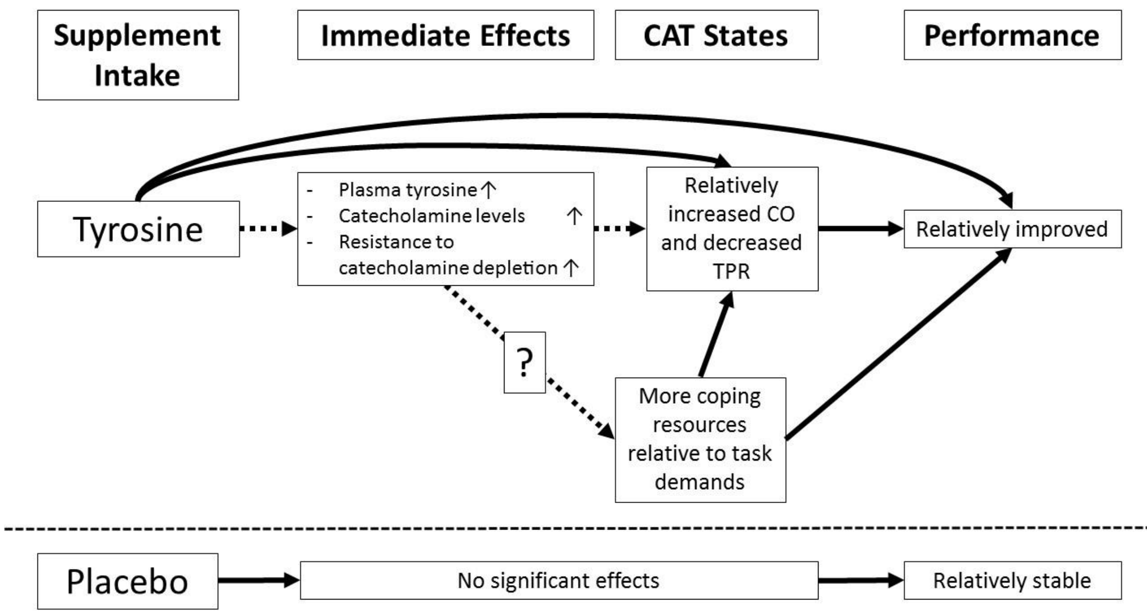

Figure 6.1. Hypothesised relationships between tyrosine intake, CAT states, and performance.

6.3 Method

6.3.1 Participants

The sample consisted of 49 students and staff members (33 male, 16 female) at a UK university, who were recruited with convenience sampling in person and through the university e-mail system. Participants were 18 to 46 years old, with a mean of 22.5 years (*SD* = 5.1). Participants' mean height and body mass were 175.0 cm (*SD* = 10.0) and 74.7 kg (*SD* = 13.6), respectively. All participants reported being healthy, right-handed or ambidextrous, and most participants were native English speakers (61%)[15]. A minimum sample size of 41 was determined with a power calculation in G*Power 3.1.9.2., using the N-Back task effect sizes (average d = 1.04) reported in Hase et al.'s (2015) systematic review, because no further effect sizes were found for the effect of tyrosine on motor performance or CAT states. Hence, the calculation used effect size d = 1.04 (f = 0.52), α = 0.05, and 90% desired power for a two-group, two-measurement comparison.

6.3.2 Materials

6.3.2.1 Cardiovascular data. The Portapres Model-2 (Finapres Medical Systems BV, Amsterdam, the Netherlands) was used to record cardiovascular variables: HR, TPR, and CO. Its measurement method is based on the arterial volume-clamp method of Peñáz (1973) and the physiological calibration criteria for the proper unloading of the finger arteries of Wesseling (1996). Further, it uses a height correction unit to compensate for hydrostatic pressure changes due to movement of the hand. It has been used in previous CAT research and allows for continuous data recording (Moore et al., 2018; Zanstra et al., 2010). It has been validated against the Finapres and the Oxford method in previous research and was found to be accurate, reliable, and cause no more missing data due to artefacts than the Oxford method (Hirschl et al., 1999; Imholz et al., 1993). Data were converted and downloaded with Beatscope version 1.1a.

6.3.2.2 Dietary supplements. Consistent with comparable previous studies (e.g., Shurtleff et al., 1994; Tumilty et al., 2014), the protocol used 150 mg / kg body mass of L-tyrosine in powder form (Myprotein.co.uk, Meridian House, Cheshire, UK) for the

[15] Native language (coded dichotomously for English versus Non-English), was not significantly correlated with performance on either of the two tasks.

tyrosine condition and 150 mg / kg body mass of microcrystalline cellulose (Blackburn Distributions Ltd, Nelson, Lancashire, UK) for the placebo condition. Both powders were mixed with 200 ml of 100% pure squeezed orange juice (Tesco Stores Ltd., Welwyn Garden City, Hertfordshire, UK).

6.3.2.3 Demand and resource evaluations. Demand and resource evaluations were assessed with two items used by previous research (e.g., Vine et al., 2013). The items were: "How demanding do you expect the upcoming task to be?" for demands and "How able are you to cope with the demands of the upcoming task?" for resources. All items were scored on a seven-point Likert scale anchored by *not at all* (1) and *extremely* (7). A cognitive CAT variable was then created from these items by subtracting demands from resources, meaning that possible scores ranged from -6 to 6 and denoted more challenge as values increased.

6.3.2.4 N-back task. The N-Back task (Kirchner, 1958) is a test of working memory that has been used in previous tyrosine supplementation research (e.g., Colzato, Jongkees, Sellaro, & Hommel, 2013). A Qualtrics survey presented a string of 23 letters (see Appendix C) for five seconds each. Starting at the fourth letter, participants were prompted to indicate (by selecting one of two boxes indicating *yes* or *no*) whether the letter shown on the current screen was the same as the letter shown three earlier (3-back condition). Thus, there were 20 items in total, 10 of them requiring *yes* and 10 of them requiring *no* as the correct answer. The maximum time was five seconds, after which the page automatically advanced if no response had been given. The number of correct answers was used as the performance outcome.

6.3.2.5 Bean-bag throwing task. Bean-bag throwing has been used as a task in previous CAT research (Turner et al., 2014). This task consisted of 20 throws of a bean-bag from a distance of 4 m to a 50x50 cm quadratic target on the laboratory floor. The bean-bag weighed 80 g and was approximately 6 cm long, 5 cm wide, and 5 cm high. Participants scored one point each time the bean-bag came to rest on the target. This scoring method was adopted in order to ensure commensurability with N-Back task scores. The number of points scored was used as the performance outcome.

6.3.3 Procedure

The study was approved by an institutional ethics committee and used a double-blind randomised crossover design. The total duration of each session was 90 minutes. One day before testing, the experimenters sent participants a list of tyrosine- or protein-rich foods to avoid in the 12 hours before testing, instructed participants not to consume any psychoactive substances (including alcohol and caffeine), and asked participants to avoid consuming any food or drinks (except water) in the last three hours before testing. Upon entering the laboratory, participants were given an information sheet and provided informed consent. The information sheet explained the study and highlighted that rewards would be given to the best three performers on each task. Participants were randomly assigned to receive either tyrosine or the placebo in the first of two testing sessions. Participants were then weighed on a SECA 770 scale (Vogel & Halke, Hamburg, Germany) in order to calculate the appropriate supplement dosage, which was mixed with orange juice by an experimenter who was not involved in the rest of the study. After consuming the drink, participants waited for 60 minutes outside of the laboratory. This wait period was consistent with findings from previous research indicating that 45-60 minutes of post-ingestion wait was sufficient to bring about significant increases in plasma tyrosine (e.g., Strüder et al., 1998; van de Rest et al., 2017). After that, a second experimenter blind to the supplement condition called participants in to sit in front of a computer, on which a Qualtrics survey was opened to guide them through the study. For the first week, participants were asked to provide demographic information and questions about their food intake on the test day before moving on to the main part of the study. The experimenter then put the Portapres on the left hand of participants, with the cuff around the middle finger and the height correction sensor around the upper arm at the height of the sternum. Participant age, sex, height, and weight were entered to calibrate the Portapres. Participants sat still for the entire duration of the cardiovascular recordings.

The order of the two tasks was randomised on each measurement occasion. Before starting each task, cardiovascular responses were recorded for a baseline of three minutes. Participants then read through the respective task instructions ($M_{\text{Reading time}}$ = 29.00 s, *SD* = 22.28 s). For each task, the survey reminded participants of the £30, £20, and £10 rewards for the best three performers, and that a quicker task completion time would determine the winner between participants with the same score. Participants then confirmed that they had read and understood the instructions. Participants were then instructed to sit still and think about the upcoming task for one minute. This minute provided the task-specific cardiovascular reactivity to be compared against the last minute of baseline. Participants subsequently completed the demand and resource evaluation items, before beginning the first task. After participants finished the first task, the procedure was repeated for the second task (baseline, task instructions, one-minute reactivity recording, demand and resource evaluation items, perform task). Approximately six minutes separated the end of the first task from the beginning of the second task. After finishing both tasks, participants were thanked for their time and reminded to return one week later at the same time to repeat the process with the other supplement.

6.3.4 Statistical Analysis

Consistent with previous research using the BPSM of CAT (e.g., Mendes et al., 2007), mean HR, TPR, and CO values were calculated for the final minute of each baseline and also for the one minute of each reactivity period. Four univariate outliers (values more extreme than three standard deviations from the mean; Stevens, 2009) were winsorised to be 1% more extreme than the next non-outlying score (adapted from Shimizu et al., 2011). The baseline values for CO and TPR were then regressed on their respective reactivity values with the standardised residuals being saved to create residualised change scores in order to adjust for baseline differences (Burt & Obradovic, 2013). TPR residualised change scores were then multiplied by -1 and summed with the CO residualised change scores to create a single physiological CAT index for each task. To test task engagement, a paired-samples t-test compared mean HR between the baseline and reactivity period.

To test the first hypothesis, paired-samples t-tests compared physiological CAT scores between the experimental conditions on each task. As an exploratory analysis, these tests were repeated for evaluations of cognitive CAT, demands, and resources. Furthermore, a correlation analysis controlling for condition examined the association between cognitive and physiological CAT scores for each task. To test the hypotheses that CAT states are associated with performance, and that performance would be better on tyrosine than on placebo, two GEE models were run to analyse the relationship between performance on each task with experimental condition, cognitive CAT, physiological CAT, and the two-way interaction terms of condition with cognitive and physiological CAT[16]. The GEE models were selected because they allow for the test of relationships between a set of independent variables and a dependent variable across different measurements, which is a parsimonious alternative to multiple separate analyses, and also allows for the inclusion of interaction effects between predictors. Significant interaction effects in the GEE analyses were probed by multiple linear regression analyses that determined simple slopes for the relationship between CAT and task performance for the respective task and condition using both CAT variables as predictors.

6.4 Results

Two participants failed to attend the second test, leading to a final sample of 47. All final analyses excluded cases that did not indicate physiological engagement with the respective task, which is a premise for the analysis of CAT states within the BPSM (Blascovich, 2008a). This lack of task engagement was evidenced by a lack of increase in HR from baseline to post-instructions[17]. For the remaining participants (37 on the N-Back task and 36 on the bean-bag throwing task), HR increased significantly from baseline to post-instructions [$M_{\text{N-Back}} = 5.34$, $SD = 3.63$, $t(53) = 10.81$, $p < .001$, $d = 1.47$; $M_{\text{Bean-bag}} = 4.79$, $SD = 3.53$, $t(44) = 9.09$, $p < .001$, $d = 1.35$]. There were no significant differences between baseline cardiovascular values for the first and second task,

---

16 In order to control for potential confounders, these analyses were repeated including age, completion time, sex, and task order as predictors. As there were no significant effects for these control variables on either task, they were not included in the main analyses. Ancillary GEE analyses also showed that they were not significantly associated with physiological CAT, although a marginally significant trend (p = 0.07) toward more challenge at older age was observed on the N-Back task.

17 On the N-Back task, 36 cases (40%) were excluded. On the bean-bag throwing task, 44 cases (49%) were excluded. Since this type of analysis has not been done before, we also report the results of our analyses using the traditional approach in an online supporting material. The significant condition effect favouring tyrosine over placebo on the bean-bag throwing task, but not the significant condition*physiological CAT interaction effect was replicated in these analyses. Though HR increased significantly on the N-Back task ($M = 1.80$, $t(89) = 2.48$, $p = .02$, $d = 0.26$), it did not significantly increase on the bean-bag throwing task ($M = 0.47$, $t(88) = 0.57$, $p = .57$, $d = 0.06$).

indicating that participants returned to their baseline values after performing ($M_{\text{Task1-Task2}}$ = -1.02; $t(44) = -0.84$, $p = .40$).

6.4.1 Comparison of CAT by Experimental Condition and Task

Table 6.1 presents descriptive statistics for systolic, diastolic, and mean arterial blood pressure; HR; CO; and TPR by task and condition. Table 6.2 summarises the paired-samples t-test comparing the placebo and tyrosine conditions on physiological CAT, cognitive CAT, demands, and resources for both tasks. There were no significant differences between conditions on the two tasks for any of the variables. Cognitive and physiological CAT were not significantly correlated on the N-Back task ($r = -.07$, $p = .61$) or the bean-bag throwing task ($r = -.10$, $p = .51$).

Table 6.1

Descriptive Statistics for Cardiovascular Data by Task and Condition

	N-Back Task								Bean-Bag Throwing Task							
	Placebo				Tyrosine				Placebo				Tyrosine			
	BL		RP		BL		RP		BL		RP		BL		RP	
	M	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
1. HR (bpm)	71.46	12.15	77.97	11.96	70.24	10.42	74.71	10.96	74.92	8.11	80.11	9.94	71.89	11.51	76.30	11.20
2. CO (lpm)	4.80	2.13	5.30	2.28	5.18	1.64	5.35	1.78	5.17	1.40	5.18	1.97	5.21	1.72	5.29	1.89
3. TPR (mmHg.s/ml)	1.30	0.62	1.29	0.88	1.16	0.50	1.13	0.58	1.13	0.49	1.20	0.61	1.12	0.61	1.24	0.80
4. SBP (mmHg)	134.33	37.23	137.67	30.88	129.74	34.28	129.22	36.11	125.40	35.18	126.22	36.18	120.17	26.40	120.76	20.66
5. DBP (mmHg)	72.58	22.02	74.87	20.18	69.52	17.86	68.84	18.27	72.12	19.63	71.95	17.79	68.49	17.14	69.85	15.31
6. MAP (mmHg)	90.28	24.83	91.90	21.80	86.61	20.45	85.48	20.98	88.16	22.75	87.60	21.89	83.88	17.75	85.14	15.73

Note. BL = Last minute of baseline period. DBP = Diastolic blood pressure. MAP = Mean arterial pressure. RP = Reactivity period. SBP = Systolic blood pressure.

Table 6.2

Descriptive Statistics and Paired-Samples T-Tests for Cognitive CAT, Physiological CAT, Demands, and Resources by Task

	N-Back Task							Bean-Bag Throwing Task						
	Placebo		Tyrosine					Placebo		Tyrosine				
	M	*SD*	*M*	*SD*	*t (df)*	*p*	*d*	*M*	*SD*	*M*	*SD*	*t (df)*	*p*	*d*
1. Cognitive CAT	1.00	2.15	0.70	1.80	1.01 (46)	.32	0.15	1.94	2.29	1.91	1.98	0.06 (46)	.95	0.01
2. Physiological CAT	0.21	2.23	-0.18	1.66	0.74 (16)	.47	0.18	0.34	1.59	0.06	1.78	0.69 (7)	.51	0.23
3. Demands	3.85	1.33	4.04	1.23	-1.01 (46)	.32	0.15	3.11	1.45	3.22	1.35	-0.48 (46)	.64	0.07
4. Resources	4.85	1.29	4.74	1.17	0.54 (46)	.59	0.08	5.05	1.44	5.13	1.27	-0.30 (46)	.77	0.04

6.4.2 Task Performance Analysis

6.4.2.1 N-back task. Table 6.3 summarises the GEE analysis of performance on the N-Back task. There were no significant main or interaction effects.

Table 6.3

GEE Parameter Estimates (N-Back Task)

Source	*B*	Wald Chi-Square	Sig.
Condition	0.55	0.77	.38
Cognitive CAT	-0.39	1.39	.24
Physiological CAT	-0.27	0.69	.41
Condition * Cognitive CAT	-0.18	0.24	.63
Condition * Physiological CAT	-0.15	0.10	.76
Intercept	15.72	814.69	.00

Note. Dependent variable: Performance. $N = 37$.

6.4.2.2 Bean-bag throwing task. Table 6.4 summarises the GEE analysis of performance on the bean-bag throwing task. There was a significant main effect for condition ($B = -1.94$, Wald $\chi^2 = 4.03$, $p = .05$, 95% CI [-3.82, -0.05]), with superior performance in the tyrosine condition than in the placebo condition. There also was a significant interaction effect for condition*physiological CAT ($B = 1.15$, Wald $\chi^2 = 5.51$, $p = .02$, 95% CI [0.19, 2.11]), with physiological CAT more positively related to performance in the placebo condition than the tyrosine condition. The additional regression analyses showed that physiological CAT was neither significantly related to performance in the placebo (*B* = 0.58, *t*[19] = 1.53, *p* = .14, sr^2 =.10), nor in the tyrosine condition (*B* = -0.58, *t*[20] = -1.76, *p* = .09, sr^2 = .13). The same was found for cognitive CAT in the placebo (*B* = 0.35, *t*[19] = 1.28, *p* = .22, sr^2 =.07) and in the tyrosine condition (*B* = 0.05, *t*[20] = 0.15, *p* = .88, sr^2 = .00).

Table 6.4

GEE Parameter Estimates (Bean-bag Throwing Task)

Source	*B*	Wald Chi-Square	Sig.
Condition	-1.94	4.03	.05
Cognitive CAT	0.05	0.04	.85
Physiological CAT	-0.58	2.23	.14
Condition * Cognitive CAT	0.30	0.68	.41
Condition * Physiological CAT	1.15	5.51	.02
Intercept	8.01	207.89	.00

Note. Dependent variable: Performance. *N* = 36.

6.5 Discussion

The present study tested whether tyrosine intake enhances challenge responses (H1) and improves performance relative to placebo on a cognitive and a motor task (H2). It also tested whether challenge responses are related to better performance than threat responses (H3). While the data did not support the first hypothesis, partial support was found for the second hypothesis as tyrosine was related to better performance than placebo on the motor task. Finally, there were no main effects for CAT states on performance, although a significant interaction effect showed that physiological CAT was more positively related to performance in the placebo condition than in the tyrosine condition.

There were no significant differences between conditions on physiological CAT. The loss of participants due to lack of task engagement may have been partially responsible for this, as small effect sizes were observed on both tasks ($d_{\text{N-Back}} = 0.18$, $d_{\text{Bean-bag}} = 0.23$; Cohen, 1992). As tyrosine has been found to be most effective in situations with high cognitive load or strong environmental stressors (Hase et al., 2015), it may be that stronger effects would be found in future studies that impose more cognitive load or stress on participants than the current study did, thereby increasing demand evaluations. This could be done by manipulating determinants of demand evaluations like uncertainty, danger, and required effort (Jones et al., 2009). The BPSM (Blascovich, 2008a) provides another potential explanation for the null findings, as it suggests that cognitive evaluations trigger physiological responses, and not vice versa. Specifically, Tomaka et al. (1997) demonstrated that evoking cardiovascular responses consistent with CAT states via exercise (versus rest) and warm (versus cold) water immersion prior to a cognitive task did not alter cognitive evaluations. As such, tyrosine might not influence cognitive evaluations. However, the BPSM acknowledges the dynamic nature of CAT states at a psychological level, for example via reappraisal. Hence, a physiological intervention that produces a noticeable effect on the psychological level might also effectively manipulate perceived coping resources and demands via reappraisal. The lack of association between the two CAT measures across both experimental conditions further complicates the conclusions drawn from the present study and poses a critical finding to the predictions of the BPSM, which posits cognitive and physiological CAT states to be interrelated (Blascovich, 2008a).

Tyrosine was associated with superior motor performance. Similarly, O'Brien et al. (2007) found that tyrosine facilitated marksmanship performance, but that effect followed cold water immersion. The current findings are thus unique in highlighting that the beneficial effect of tyrosine on motor performance is not contingent on cold water immersion. The lack of significant differences between tyrosine and placebo on the present cognitive task is inconsistent with previous findings from studies with and without cold exposure (Colzato et al., 2013; Mahoney et al., 2007; O'Brien et al., 2007). However, only one of these studies used the N-Back task (Colzato et al., 2013). Although that study found significant differences between tyrosine and placebo on a less demanding condition of the N-Back task (2-Back), it featured a greater number of stimuli, shorter presentation time per stimulus, and shorter stimulus-onset asynchrony. It is unclear whether these differences caused participants to perceive higher demands and feel more pressurised. An alternative explanation could be that the 2-back condition simplified the working memory component of the task enough to let other domains of cognitive function become the deciding factor in determining performance (e.g., sustained attention or response execution rather than working memory). This could serve to explain why different results were found in the past and present studies.

On the motor task, there was a significant interaction effect between condition and physiological CAT. In particular, physiological CAT was more positively related to performance in the placebo condition than in the tyrosine condition. Follow-up analyses revealed that although the regression slope for physiological CAT was in the predicted direction in the placebo condition, this trend was not statistically significant. In the tyrosine condition, the trend was in the opposite direction. This finding is inconsistent with the general predictions of the BPSM (Blascovich, 2008a) and the findings of a recent systematic review of the relationship between CAT states and performance (chapter 2). They might in part be explained by the temporal gap between CAT measurement and task performance, allowing for variation in CAT states, although previous research has found a relationship between CAT states and performance with comparable or even longer gaps (e.g., Blascovich et al., 2004). Similarly, the relatively large number of trials could also have provoked variation in CAT states throughout task performance, therefore attenuating the relationship between the initial CAT measurement and performance at the end of the task. The fact that the relationship between physiological CAT and performance in the tyrosine condition was negative (albeit non-significantly so) might appear counterintuitive, but could suggest that tyrosine is particularly beneficial for those individuals experiencing a threat state and less helpful for those in a challenge state, potentially even hampering performance for strongly challenged individuals.

Given the lack of differences between conditions on the CAT variables in the present study, alternative pathways through which tyrosine exerts beneficial effects on performance warrant consideration. Rather than directly influencing CAT states, the current findings suggest that tyrosine may operate independently to improve motor performance. Although this independent mechanism has not been explored yet, a possible candidate could be an effect of tyrosine on dopamine function in the striatum, whose activation has been linked with areas associated with action preparation and execution, such as the postcentral gyrus, precentral gyrus, and supplementary motor area (Molenberghs, Trautwein, Böckler, Singer, & Kanske, 2016). However, future research should examine whether this finding can be replicated and explained in more detail. For example, research could identify whether tyrosine helps threatened individuals to actually adopt a challenge state while performing a task, or whether these individuals remain threatened, but still outperform challenged individuals.

Despite the strengths of the study in exploring the impact of a dietary supplement on CAT states and performance across both a cognitive and motor task, some limitations should be acknowledged. Although participants were encouraged to perform well and financial incentives were offered, task engagement was still low in some participants. Specifically, some participants showed decreases or no change in HR, failing to meet the BPSM's premise of task engagement (Blascovich, 2008a), and were subsequently excluded from the analyses. The lack of verbally delivered instructions and extrinsic motivators such as performance-contingent punishments and social evaluation might be partly responsible for this. Further, the mean increases in HR were rather small, although it should be noted that during the recordings, participants were seated and quietly imagined the upcoming task, which should provoke lesser increases in HR due to being less metabolically demanding than, for example, holding a speech (e.g., Blascovich et al., 2004). The lack of a VC measure also limits the study, as an index based on HR and VC could have been a more robust indicator of task engagement than HR reactivity alone (e.g., Streamer et al., 2017).

Another limitation concerns the generalisability of the findings to well-learned tasks or metabolically demanding tasks (i.e., anaerobic performance; Jones et al., 2009), as both tasks in the present study were novel to the vast majority of participants and did not involve any strenuous physical exercise. A field study in a high-pressure environment (e.g., a professional sports competition) could prevent these limitations by examining expert performance in participants likely to show greater task engagement. A third limitation is the lack of a manipulation check comparing plasma tyrosine and catecholamine levels immediately before supplement ingestion and testing. However, similarly designed studies that used an equal or slightly lower dosage have found that plasma tyrosine increased significantly within 60 minutes of consumption (Strüder et al., 1998; Tumilty et al., 2014), and that tyrosine may increase plasma catecholamines relative to placebo (Kishore et al., 2013).

Future research could measure physiological CAT states throughout task performance in order to explore the dynamic relationship between CAT states and performance and the present finding that tyrosine can benefit individuals in a threat state more than those in a challenge state. More specifically, research could test whether the negative relationship between CAT states and performance on tyrosine will persist during task performance, or whether it promotes a challenge state in threatened participants during task performance, but not during task preparation. Future work could also benefit from increasing the ecological validity of tyrosine supplementation research. For example, it would be important to know whether the present findings generalise to contexts of sports competitions or university exams, and to non-fasted participants with varying dietary habits. Indeed, the relationship between CAT states and performance has been explored in sports competitions and university exams, but studies have yet to examine the impact of tyrosine intake on CAT states in those contexts (Blascovich et al., 2004; Seery et al., 2010). Further, research on CAT manipulations is still limited. With the current exception, research has only manipulated psychological antecedents of CAT states with instructional sets or other psychological techniques (e.g., Feinberg & Aiello, 2010; Moore et al., 2015). The BPSM of CAT provides other possibilities for physiological CAT interventions that warrant exploration (e.g., decreasing TPR with the nitric oxide precursor L-arginine; Moncada, Palmer, & Higgs, 1991). Ultimately, sports psychologists and other professionals should look to develop a multi-method toolkit containing several interventions that can reliably promote a challenge state or buffer the detrimental effect of a threat state on performance.

6.5.1 Conclusion

The present study was the first to test the effects of tyrosine intake relative to placebo in a BPSM framework. In a financially incentivised competitive setting, tyrosine was associated with better performance than placebo on a motor task. Tyrosine produced no significant differences on cognitive evaluations and cardiovascular responses. However, cardiovascular responses were negatively related to performance on tyrosine, while a positive trend was found on placebo. The finding that tyrosine improved motor performance holds relevance for individuals requiring fine motor performance, as tyrosine presents an effective and safe supplement to optimise their performance under pressure.

Table 6.5

Summary of Chapter 6 and Preview of Next Chapter

Chapter	Aim	Findings
6	To examine whether tyrosine intake promotes a challenge state.	Tyrosine did not significantly affect cognitive or cardiovascular CAT responses, although small effect sizes were observed on cardiovascular CAT. The relationship between cardiovascular CAT and performance was significantly less positive in the tyrosine than in the placebo condition, suggesting that tyrosine may especially benefit participants in a threat state.
	Rationale for next chapter The meta-analysis in chapter 7 did not follow logically from chapter 6, but rather from the evaluation of the data collected for the entire thesis. Precisely, I reasoned that the data collected throughout the thesis (unlike the data incorporated in chapter 2) were homogenous enough to permit a meta-analysis of the relationship between cognitive and cardiovascular CAT variables. This would in turn provide more robust conclusions drawn from this thesis regarding the cognitive-cardiovascular CAT relationship.	
Chapter	Aim	Findings
7	To conduct a meta-analysis of the relationship between cognitive and cardiovascular indicators of CAT states collected in all empirical studies of this thesis.	

Chapter 7

Internal Meta-Analysis of Cognitive-Cardiovascular Challenge and Threat Relationship

7.1 Abstract

Although the biopsychosocial model of challenge and threat specifies that cognitive evaluations and cardiovascular responses reflective of challenge and threat states should be positively related, no strong consensus has emerged from the literature. Since the studies conducted for this thesis have consistently collected cognitive and cardiovascular challenge and threat measures, this chapter aimed to examine the relationship between the two measures in the present data. To do this, a brief meta-analysis of the relationship between cognitive demands-resources difference scores and cardiovascular challenge and threat indices was conducted across the four empirical studies of this dissertation. Two versions of the meta-analysis are reported: One including imagination-based, and one including speech-based cardiovascular challenge and threat data from chapter 4. In both versions, 21 associations were analysed and the meta-analysis indicated a small but significant negative association between cognitive challenge and threat evaluations and cardiovascular challenge and threat responses [version 1: g = -.10, p = .01, 95% CI (-.18, -.02); version 2: g = -.10, p = .01, 95% CI (-.18, -.02)]. Thus, this thesis found no consistent support for the predictions of the biopsychosocial model and even implied a negative association. Future research should examine whether this finding was due to moderating influences, such as biased responding to self-report items of cognitive evaluations (e.g., by including a social desirability scale) and/or the timing of cardiovascular measurements (i.e., during or before task), or whether it reflects an incorrect specification of the biopsychosocial model that should be corrected by future theoretical frameworks.

7.2 Introduction

Research using a BPSM perspective to study CAT states in motivated performance situations has advanced the field in many ways by answering existing research questions creatively and constructively. However, one question that has not been definitively answered is whether the prediction of the BPSM that cognitive CAT evaluations and cardiovascular CAT responses are positively related can withstand empirical scrutiny. Indeed, one of the most common study limitations in the published CAT literature is that only one measure of CAT states was collected, or that no association between the two measures was reported (see chapter 2). The few studies that reported a correlation provided equivocal evidence (e.g., Rith-Najarian et al., 2014; Turner et al., 2013; Vine et al., 2013). Though some studies reported a positive association (e.g., Vine et al., 2013; Zanstra et al., 2010), the effect sizes were small (e.g., Moore et al., 2018) to moderate (Vine et al., 2013). Thus, the field would benefit from a quantitative synthesis of the observed effects to produce more robust evidence supporting or questioning the prediction of the BSPM. Meta-analysis is a suitable method for such a quantitative synthesis (Glass, 1976; Schulze, 2004). As this thesis reports on cognitive and cardiovascular CAT states data from four separate studies comprising 21 measurement occasions, a meta-analysis seemed useful to examine the general association between cognitive CAT evaluations and cardiovascular CAT responses. Unlike the considerably heterogeneous evidence synthesised by chapter 2's systematic review, the data collected for this thesis (and the cognitive-cardiovascular CAT comparison in particular) were homogenous enough to warrant a meta-analysis. For example, the methods of measuring demand and resource evaluations and constructing the single cognitive and cardiovascular CAT variables were the same across the different studies reported in this thesis (the second cognitive CAT score from chapter 4 was not considered here as it was only used for the variance components analyses). This contrasts the evidence included in chapter 2 (see section 2.4.3.1), which was characterised by various methods of producing single cognitive and cardiovascular CAT variables. Thus, the data collected for this thesis permit a meta-analytic examination of the relationship between cognitive and cardiovascular CAT scores. Such a meta-analysis could improve the conclusions about the evidence for or against a positive association between cognitive and cardiovascular CAT states. The BPSM provided the main hypothesis for this meta-analysis, being that cognitive and cardiovascular indicators of CAT states are positively associated.

7.3 Method

7.3.1 Participants

Data from the four empirical studies of this dissertation comprising 171 participants (M_{Age} = 21.6 years, *SD* = 6.2, $\%_{\text{female}}$ = 26%) were used. Table 7.1 details sample characteristics of the four studies and the overall meta-analysis.

Table 7.1

Sample Characteristics across Studies

Chapter	Participants	N	M_{Age}	SD_{Age}	Age range	N_{Female}	N_{Male}
3	University students and staff members	30	23.4	4.9	18-35	2	28
4	Elite trampoline athletes	30	14.6	3.4	10-22	17	13
5	University students and staff members	62	23.5	6.3	18-52	10	52
6	University students and staff members	49	22.5	5.1	18-46	16	33

Total	171	21.6	6.2	10-52	45	126

7.3.2 Materials

7.3.2.1 Cardiovascular data. See chapters 3-6 for cardiovascular equipment and data collection methods. The meta-analysis used the residualised change score-based cardiovascular CAT index computed and used in each empirical study chapter.

7.3.2.2 Demand and resource evaluations. See chapters 3-6 for demand and resource evaluation materials and methods. The meta-analysis used the resources-demands difference score described and used in each empirical study chapter.

7.3.3 Statistical Analysis

Pearson correlation coefficients were computed for each measurement during which both a cognitive CAT and a cardiovascular CAT variable had been collected. In chapter 4, the cognitive CAT variable was separately correlated with the imagination-based (further referred to as chapter 4a) and the speech-based cardiovascular CAT indices (further referred to as chapter 4b). The metacor package (Laliberté, 2011) was used to meta-analyse the correlations in RStudio, version 1.0.143. Mean correlation coefficients, 95% confidence intervals, and p-values were computed with the metacor.OP function, which uses the Olkin-Pratt method (Olkin & Pratt, 1958). Effect sizes were classified according to Cohen (1992). To examine whether the associations observed in this thesis were biased toward significant results, the Meta-essentials package (Suurmond, van Rhee, & Hak, 2017) was used to provide funnel plots and fail-safe N statistics according to Rosenthal (1983). Rosenthal's ad-hoc rule was used to determine whether a small or large number of studies with an average effect of zero would need to exist to render the results of this meta-analysis non-significant. The Meta-essentials package (Suurmond et al., 2017) was also used to provide heterogeneity statistics (Cochran's *Q*, I^2).

7.4 Results

The four empirical studies of this dissertation provided 24 Pearson correlation coefficients from 613 observations in total. Table 7.2 details the correlations between cognitive and cardiovascular CAT in the individual studies. As there were two cardiovascular CAT variables, but only one cognitive CAT variable for each of the three measurements in chapter 4, the analysis was conducted once with the imagination-based (version 1; see chapter 4a in table 7.2), and once with the speech-based cardiovascular CAT variable (version 2; see chapter 4b in table 7.2) for each measurement. Thus, the final analyses aggregated a total of 21 correlation coefficients each. The mean g statistic was -.10 for version 1 [p = .01, 95% CI (-.18, -.02)] and -.10 for version 2 [p = .01, 95% CI (-.18, -.02)]. These associations qualified as small effect sizes (Cohen, 1992). There was insufficient evidence to infer heterogeneity in version 1 (Q = 13.87, p = .84, I^2 = 0.00%) or version 2 (Q = 13.43, p = .86, I^2 = 0.00%). The fail-safe N statistics according to Rosenthal (1979) were 22 (version 1) and 21 (version 2). According to Rosenthal's ad-hoc rule, the statistics indicated that a small number of studies averaging a *Z*-value of zero would be required to make the combined effect size statistically non-significant. Figures 7.1 and 7.2 present funnel plots for meta-analysis versions 1 and 2, respectively.

Table 3

Cognitive-Cardiovascular CAT Correlations by Study and Measurement

Chapter	Analysis	Note	*r*	*n*
3	1	W1, SUT	-.09	29
	2	W1, NBT	.01	26
	3	W1, BBT	-.15	26
	4	W1, DTT	.29	29
	5	W2, SUT	-.21	26
	6	W2, NBT	-.27	26
	7	W2, BBT	-.07	23
	8	W2, DTT	-.23	25
	9	W3, SUT	.25	29
	10	W3, NBT	-.04	30
	11	W3, BBT	-.23	29
	12	W3, DTT	-.33	30
4a – Imagination-based	1	M1	-.20	25
	2	M2	-.07	27
	3	M3	-.26	28
4b – Speech-based	1	M1	-.13	25
	2	M2	-.17	27
	3	M3	-.21	28
5	1	BL	.01	54
	2	FT	-.05	52
6	1	PLA, NBT	-.01	23
	2	TYR, NBT	-.22	31
	3	PLA, BBT	-.06	22
	4	TYR, BBT	-.22	23
Version 1		Excluding chapter 4b data	-.10	613
Version 2		Excluding chapter 4a data	-.10	613

Note. W = Week. SUT = Subtraction task. NBT = N-Back task. BBT = Bean-bag throwing task. DTT = Dart-throwing task. M = Measurement. PLA = Placebo. TYR = Tyrosine.

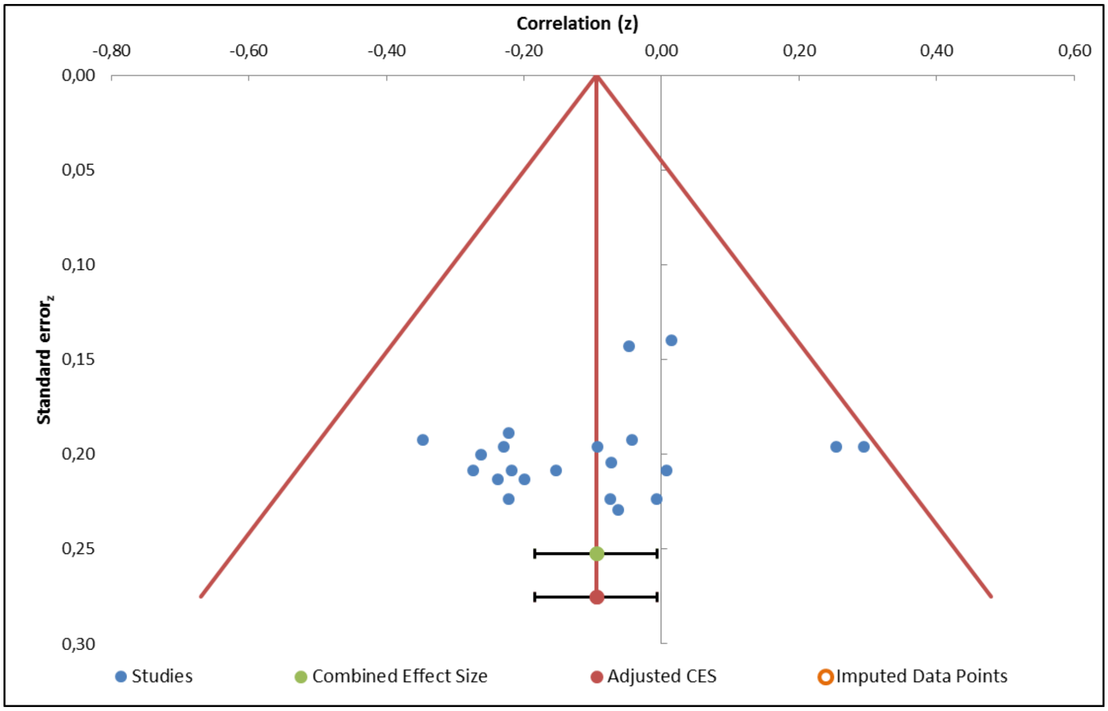

Figure 7.1. Funnel plot for meta-analysis, version 1.

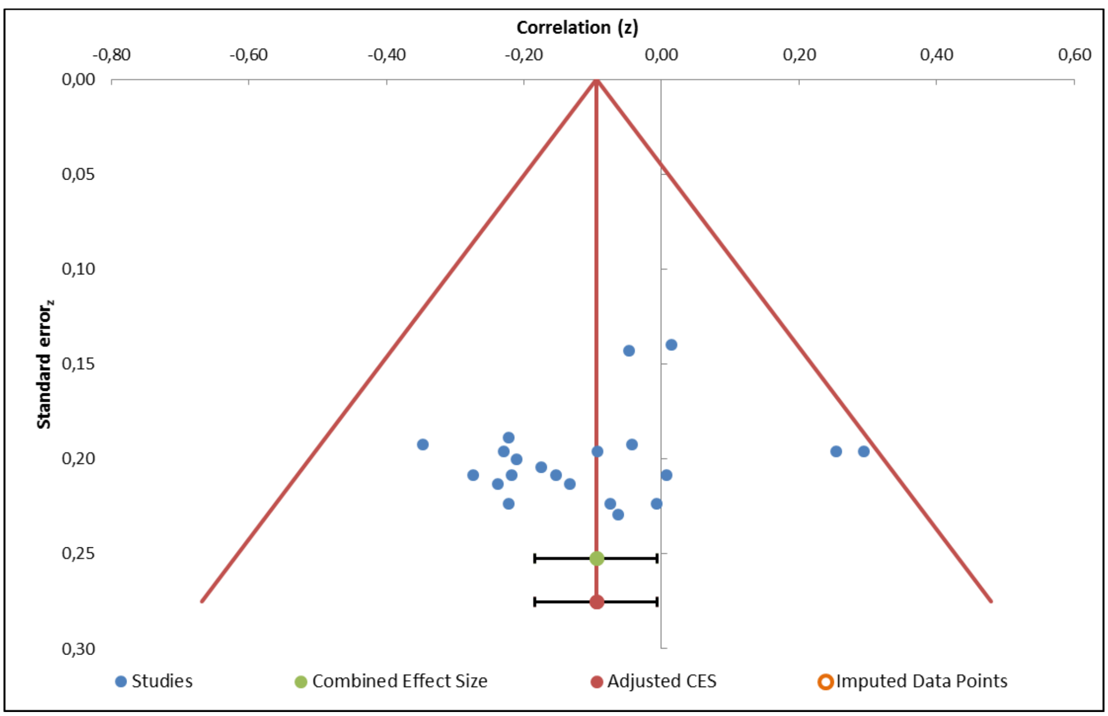

Figure 7.2. Funnel plot for meta-analysis, version 2.

7.5 Discussion

The brief meta-analysis showed that in the research conducted for this thesis, cognitive CAT evaluations were generally negatively related to cardiovascular CAT responses. The average relationship amounted to a small effect that stands in contrast to the predictions of the BPSM (Blascovich, 2008a) and some previous empirical studies explicitly testing and finding support for the prediction that cognitive CAT evaluations are positively related to cardiovascular CAT responses (e.g., Vine et al., 2013; Zanstra et al., 2010). The meta-analysis also showed that while the observed associations did not seem to be biased toward significant results and there was insufficient evidence to infer heterogeneity, a small number of null-effect studies would be required to render the overall effects non-significant.

The finding that cognitive CAT evaluations were negatively associated with cardiovascular CAT responses in this thesis and not consistently positively associated in the literature poses an important question for the BPSM. Naturally, even if cognitive CAT evaluations and cardiovascular CAT responses independently predict performance (see chapter 2), the lack of a positive association between them leads to the question of what they actually represent. As the BPSM specifies cognitive evaluations and cardiovascular responses to be part of the same underlying stress response, this would be an interesting question to address in future research. However, next to the possibility of two distinct processes in the stress response operating independently, there is another possibility.

As Blascovich and Mendes (2000) already pointed out, some individuals may not be conscious of the cognitive evaluation process, or may not be able to accurately assess their demands and resources, which might be due to insufficient cognitive abilities or cognitive biases. Thus, cognitive CAT evaluations might be positively associated with cardiovascular CAT responses when controlling for cognitive abilities and biases. This also provokes the question of whether interventions can be implemented to aid the accurate assessment of cognitive CAT evaluations. For example, one could try providing more detailed instructions for the self-report items to make it clearer to individuals as to what they should be assessing and reporting. Also, one could examine whether an (at this point entirely hypothetical) implicit cognitive CAT measure might more accurately measure cognitive CAT evaluations. A final question concerns the timing of the cardiovascular responses. Although Blascovich and Mendes (2000) considered physiological CAT measures superior to self-report measures for not being prone to psychological biases, there still is potential for inaccuracy in physiological CAT measures. Precisely, it could be that the measurement of cardiovascular responses during, rather than prior to, performance provide a stronger relationship between cardiovascular and cognitive indicators of CAT states, as well as with performance. However, cardiovascular responses recorded during task performance would have been confounded by movement on most of the tasks employed in this thesis (e.g., trampoline jumping, bean-bag throwing, dart-throwing).

Even though a small effect was observed that suggests a negative association between cognitive CAT evaluations and cardiovascular CAT responses in this thesis, the fail-safe N statistics indicated that a small number of null effects would be required to render the overall effects observed in this meta-analysis non-significant. Thus, further examination of the association between cognitive CAT evaluations and cardiovascular CAT responses across the literature is warranted to provide more robust conclusions.

7.5.1 Conclusion

A meta-analysis found that the associations between cognitive CAT evaluations and cardiovascular CAT responses were negative on average, with small effects being observed. This finding is at odds with the predictions of the BPSM and thus raises the question of whether cognitive and cardiovascular indicators of CAT states actually represent distinct phenomena or whether the measurement of cognitive, cardiovascular, or both indicators of CAT states require greater precision to replicate the associations specified by the BPSM and found by select studies.

Table 7.3

Summary of Chapter 7 and Preview of Next Chapter

Chapter	Aim	Findings
7	To conduct a meta-analysis of the relationship between cognitive and cardiovascular indicators of CAT states collected in all empirical studies of this thesis.	The meta-analysis produced a significantly negative average correlation of small effect size between cognitive and cardiovascular indicators of CAT states.
8	To discuss main findings, significance and implications, limitations, and future research directions of this thesis.	

Chapter 8

General Discussion

8.1 Main Findings

This thesis had the following aims: a) to systematically review the relationship between CAT states and performance; b) to examine the relative contributions of personal, situational, and person by situation interactional factors in CAT states; c) to examine the potential of selected interventions to promote a challenge state or mitigate the detrimental effects of a threat state; and d) to examine the interrelationships between cognitive and cardiovascular indicators of CAT states and performance in original empirical research. Figure 8.1 graphically summarises the research conducted to address these aims; providing method, findings, and contribution of findings by chapter.

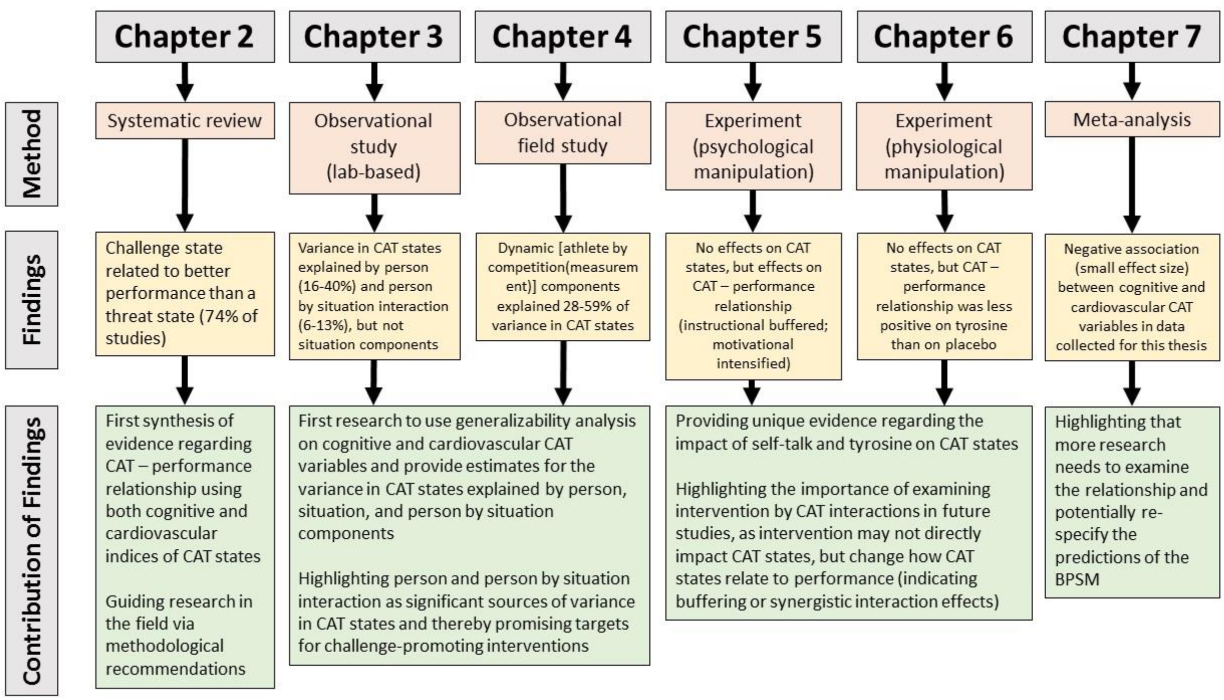

Figure 8.1. Method, findings, and contribution of findings by chapter.

In a systematic literature review of 38 published studies conceptualising CAT states consistent with the BPSM, the majority of effects (74%) indicated that a challenge state was associated with better performance than a threat state. There was significant variation in the reviewed studies regarding what CAT measures, outcome tasks, and research designs they utilised. The benefits of a challenge state on performance were largely consistent across studies using cognitive, physiological, and dichotomous CAT measures, cognitive and behavioural outcome tasks, and direct experimental, indirect experimental, correlational, and quasi-experimental research designs. When putting these findings into perspective, it is noteworthy that there were no previous systematic reviews or meta-analyses of the associations between CAT states and performance in a BPSM framework. However, these findings converge with the findings of a meta-analysis that was conducted concomitantly and independently (Behnke & Kaczmarek, 2018). This meta-analysis supported the superiority of a challenge over a threat state on various performance outcomes and research designs, although it only looked at cardiovascular markers of CAT states and highlighted a risk of publication bias in the published studies. Furthermore, the results of the present systematic review are also consistent with the predictions of the BPSM and the TCTSA.

Chapters 3 and 4 aimed to decompose the variance in cognitive and cardiovascular measures of CAT states measured in a laboratory-based, fully-crossed (chapter 3) and a field-based, nested (chapter 4) study design. Variance components analyses tested the variance explained by differences between participants (person component) and differences in situational or dynamic factors. In chapter 3, situational factors were split into differences between tasks and time points (situational components), and interactions thereof with the person component (person by situation component). Due to the nested design in chapter 4, the situational and person by situation interactional components were reduced to a dynamic component [competition(athlete)] in that chapter. The person component explained a significant proportion of the variance in CAT states on most outcomes in chapter 3, but only on resource evaluations in chapter 4. In chapter 3, the situation component did not explain a significant proportion of variance. Significant person by situation interaction components were found on cognitive CAT evaluations in chapter 3. The dynamic component of chapter 4 that represented situational and person by situation interactional factors explained a significant proportion of variance on cognitive and cardiovascular CAT outcomes.

As no previous publication has reported on variance components analyses of CAT states, there is no direct evidence to put these findings into perspective. A study by Lucas and colleagues (2012) provided evidence somewhat comparable to the present findings on cognitive CAT evaluations, as they collected self-reports of primary and secondary stress appraisals from police officers who rated the hypothetical stressfulness of different job scenarios. The study found person, situation, and person by situation interaction components, which is partly consistent with the present findings in that person and person by situation interaction components were found. However, they are partly inconsistent in that no situation component was found. This might be in part due to statistical power issues, as the task component explained 6-12% of the variance on cognitive CAT evaluations in chapter 3. However, the lack of significant variance components in the remaining analyses also poses the question of whether other situational variables (e.g., evaluative observation versus performing alone, participating at a certain time of day/day of week) might have been needed to consistently explain significant amounts of the variance in CAT states. A further inconsistency is that the percentages attributed to each variance component differ. For example, whereas the person component in police officers explained 14-15%, it varied much more in the present studies (0-40%). Conversely, whereas the person by situation component explained 38-41% in police officers, it explained less variance in chapter 3 (17-22%). The corresponding component in chapter 4 [competition(athlete)] explained a greater amount of variance (25-59%). There were some key features of the present work that represent advancements from Lucas and colleagues' work. For example, the present thesis reports not only on self-reports of cognitive evaluations (the equivalent to Lucas and colleagues' stress appraisal measures), but also on cardiovascular CAT variables. Moreover, where Lucas and colleagues asked officers to provide general ratings for non-specific stressors (e.g., "inadequate supervisor support", "demand for high morality", etc.), the present research collected data on-site as participants prepared for their specific upcoming performances (chapter 4).

Two experimental studies tested the effects of two psychological interventions (instructional and motivational self-talk; chapter 5) and one physiological intervention (a tyrosine supplement, chapter 6) that had not previously been examined in CAT research. Although none of the interventions affected CAT states directly (i.e., promoted a challenge state), two of them (tyrosine and instructional self-talk) impacted how cardiovascular CAT responses related to performance in that a threat state was relatively less detrimental than in the placebo/control condition. However, motivational self-talk intensified the relationship between cardiovascular CAT responses and performance. As no previous studies have examined the impact of self-talk or tyrosine interventions on CAT states, there are no directly conflicting results. However, even other studies that have administered interventions to manipulate CAT states have not examined any interaction effects between interventions and continuous measures of CAT states in predicting performance. For example, two studies successfully promoted a challenge or mitigated a threat state with an arousal reappraisal intervention, but did not report on the potentially differential relationships between continuous CAT measures and performance within the experimental groups (Moore et al., 2015; Sammy et al., 2017). Another study did report an interaction effect between CAT and task structure, where participants in a challenge condition only reached better negotiation outcomes than those in the threat condition when the negotiation task had integrative potential (versus being purely distributive; O'Connor et al., 2010). However, as this effect represented the interaction of two dichotomous variables (both created with a manipulation), it cannot be interpreted in the same way as the interaction effects observed in this thesis.

The data collected for this thesis also allowed for the examination of the relationships between cognitive and cardiovascular indicators of CAT states and performance. The relationship between CAT states and performance was not consistently supported, as only chapter 5 found cognitive evaluations consistent with a challenge state to be related to significantly better performance than those consistent with a threat state, and only chapter 4 found one of its (speech-based) cardiovascular CAT variables to be related to performance in the same way. Moreover, chapter 4 found relationships between cognitive evaluations and performance that went in the opposite direction, which may have been due to the young age of the sample (e.g., Rith-Najarian et al., 2014; cf. Turner et al., 2013). These results are inconsistent with the predictions of the BPSM and the TCTSA, and findings of previous research (as reviewed in chapter 2; see also Behnke & Kaczmarek, 2018) that generally found a challenge state to relate to better performance than a threat state both in terms of cognitive evaluations and cardiovascular responses. Furthermore, the relationship between cognitive and cardiovascular indicators of CAT states as observed throughout the empirical studies of this thesis was inconsistent with the BPSM. A meta-analytic aggregation of all relationships between the single cognitive CAT score and the cardiovascular CAT index used throughout this thesis indicated that the average relationship between the two variables was negative and of small effect size (Cohen, 1992). This indicates that the BPSM might need to be more clearly specified regarding the relationships between cognitive and cardiovascular CAT measures. Although Blascovich and Mendes (2000) already highlighted that self-reports of cognitive evaluations are more problematic than cardiovascular measurements due to biases or difficulties in the evaluation process, they did not specify any specific methods to treat these biases or difficulties.

8.2 Significance and Implications

The research conducted for this thesis has important implications for researchers who use the BPSM or other CAT theoretical models, and for applied sport psychologists interested in improving sport performance. First, the systematic review of the relationship between CAT states and performance (chapter 2) presented an important milestone to the CAT literature, as no previous article had analysed and summarised the relationship between CAT states and performance across CAT measures, outcome tasks, and research designs. Indeed, the publication of a meta-analysis of the relationship between cardiovascular CAT variables and performance around the same time by an independent research group showed that there was a need for such an analysis of the previously published studies (Behnke & Kaczmarek, 2018). Thus, chapter 2 addressed a need from the field and supported the predictions of the BPSM and TCTSA in a way that advanced the evidence base. It also reaffirmed researchers and applied sport psychologists alike in researching and implementing reliable CAT monitoring systems and challenge-promoting interventions in pursuit of peak performance. Furthermore, chapters 1 and 2 highlighted some gaps in the literature and thereby laid the groundwork for the empirical studies of the present project. For instance, chapters 3 and 4 addressed the gap regarding the relationship between personal and situational components of CAT states that was highlighted in chapter 1. They also presented the first repeated-measures studies in the CAT literature to have measured CAT states over several weeks. Chapter 4 also addressed the limitation of low ecological validity in chapter 3 by sampling elite-level athletes performing at real-world competitions. Finally, chapters 5 and 6 addressed the calls for novel intervention studies in chapters 1 and 2.

Second, the finding that CAT states appear to vary largely as a function of person and person by situation interaction components in a laboratory-based context (chapter 3), and as a function of dynamic components in a field study-based motivated performance context (chapter 4) implies that researchers and practitioners should consider individual difference variables, and how they interact with certain situations to produce a challenge or a threat state. For example, challenge-promoting interventions could target individual difference variables that play a role in CAT states. Chapter 5 examined self-talk as a variable to be potentially associated with CAT states. Although self-talk is commonly viewed as a strategically used intervention, self-talk use also occurs naturally and likely varies between individuals (Hardy et al., 2005; Latinjak, Hatzigeorgiadis et al., 2019). Chapter 6 administered tyrosine, which could be used habitually to target individual differences in catecholamine function (e.g., Cools, 2006; Jongkees, Hommel, & Colzato, 2014), which might in turn reflect part of the person component in CAT states.

Researchers have already begun to consider individual differences in CAT evaluations. Tomaka and colleagues produced a questionnaire assessing stable tendencies regarding whether individuals habitually react with a challenge or a threat state across situations (Tomaka et al., 2018). This research indicated that stable tendencies to respond to a stressful situation with a challenge state (versus a threat state) were associated with better mental health status (e.g., less depression and post-traumatic stress disorder symptoms, greater life satisfaction). Such a questionnaire could be used to predict CAT responses across situations, but also to associate the stable CAT disposition with other variables. For example, the stable CAT disposition could be associated with self-efficacy and achievement goal orientation, which according to the TCTSA function as antecedents to CAT states and have both been shown to at least partly reflect stable tendencies across different situations (Smith, 1989; Bandura, Caprara, Barbaranelli, Gerbino, & Pastorelli, 2003; Muis & Edwards, 2009). Furthermore, a stable CAT disposition could be associated with physiological variables, such as hormone levels and their reactivity to stressors. An example could be testosterone, which has been shown to be highly variable between individuals in response to motivated performance situations and was suggested to be associated with cognitive appraisal (Oliveira & Oliveira, 2014; Salvador & Costa, 2009). Furthermore, testosterone has been positively associated with nitric oxide synthesis, which in turn elicits vasodilation and could therefore explain a testosterone-based challenge cardiovascular response (Goglia et al., 2010).

Third, the findings from the experimental studies presented in this thesis also bear relevance for researchers and practitioners. For example, the finding that instructional self-talk and tyrosine attenuated the relationship between cardiovascular CAT states and performance carries important implications for applied sport psychologists. For example, the finding implies that although not every athlete might be able to reach a challenge state before every competition, negative performance consequences in threatened athletes might still be averted with the help of interventions such as instructional self-talk or tyrosine intake. This would be in line with other intervention research that focused on mitigating a threat state (Moore et al., 2015). Conversely, motivational self-talk could be used as a performance enhancer for athletes already in a challenge state. Psychological researchers could also benefit from these findings by acknowledging the possibility of absent relationships between CAT states and performance in the entire sample, but not in specific subgroups. Thus, the present findings could stimulate greater awareness of intervention-related moderator effects on the relationship between CAT states and performance (as well as moderator effects in observational studies).

Fourth, the examination of the interrelationships between cognitive and cardiovascular indicators of CAT states and performance throughout this thesis produced some significant implications for the BPSM, as the present results were inconsistent with the predictions of the BPSM. This thesis revealed that CAT researchers should be aware of potential moderators of the relationship between CAT states and performance, such as age of their participants (which may moderate the relationship between cognitive CAT and performance) or interventions that they intend to administer (which may have a synergistic or buffering effect on the relationship between cardiovascular CAT and performance). Also, the relationship between cognitive CAT evaluations and cardiovascular CAT responses should be explored further, as it might be significantly influenced by other moderators. For example, the deliberations of Blascovich and Mendes (2000) regarding the ability to accurately assess situational demands and coping resources could be used as a starting point for research into moderators of the relationship. However, the negative association observed between cognitive and cardiovascular CAT measures in this thesis (chapter 7) combined with the lack of consistent support for the association in other studies following the publication of the

BPSM might also imply that the model requires to be revised regarding this association. This thesis has contributed to the field by indicating that the predictions of the BPSM may not always hold in a sport psychological research context and by showing that this may be due to moderators acting on the relationship between CAT states and performance.

Last, this thesis holds some methodological implications. First of all, the finding of sustained task engagement across repetitions of the same motivated performance situation (chapter 3) encourages researchers looking to examine CAT states across repeated measurements. Although one should continue checking task engagement at each new measurement, the present results indicate that CAT research taking measurements in different weeks may not suffer from time-related decreases in task engagement. However, other research indicates that more care should be taken when employing the same task multiple times in the same testing session (Kelsey et al., 2004). Another methodological implication comes from the direct comparison of an imagination-based and a speech-based cardiovascular CAT variable in chapter 4. The speech-based variable exhibited more positive relationships with performance and greater HR reactivity, although it is not clear whether the latter finding purely reflected greater task engagement or just the heightened metabolic demands of speech production. Certainly, these results appear to favour speech-based cardiovascular CAT indices over imagination-based ones. A final methodological implication concerns the statistical analysis of the relationship between CAT variables and performance in intervention studies, which may be attenuated or intensified by the interventions. Thus, I recommend including intervention by CAT variable interaction effects to assess potential effects of interventions on the relationship between CAT states and performance. Also, when testing the relationship between CAT states and performance as specified by the BPSM or another theory, then researchers should examine this prediction in appropriate subgroups only. For example, a placebo or control group would be more appropriate for examining this relationship than an intervention group that was effectively manipulated into a challenge or a threat state (due to the reduced diversity in CAT states in the intervention group).

8.3 Limitations

This thesis has some limitations that should be taken into account when interpreting its findings. For example, the BPSM specifies that the pituitary-adrenocortical activation in a threat state involves cortisol release, which then offsets the effects of the sympathetic-adrenomedullary activation that is given in both CAT states. However, no cortisol or other hormonal measures were collected in the present research. Though the cardiovascular indicators of CAT states are widely accepted (Blascovich, 2008a; Seery, 2013) as sufficient to denote physiological differences between CAT states, collecting hormonal measures could have enriched the knowledge generated. The research in this thesis was also limited by the lack of a VC measure, which would have allowed for a more robust inference of task engagement than using HR only.

Furthermore, extrinsic motivators were used to a lesser extent than in previous work (e.g., Moore et al., 2012; Sammy et al., 2017). For example, no scoreboard with fictitious results was displayed to participants, no video recordings were staged, and no interview following poor performance was announced in order to increase pressure (e.g., Moore et al., 2012). Doing this might have prevented the minor task engagement issues in chapters 5 and 6, but generally participants exhibited sufficient task engagement throughout the studies presented in this thesis.

Finally, some of the findings of this thesis may be limited by low ecological validity. Whereas chapters 3 and 4 used similar methodologies in different samples to ensure the generalisability to both athlete and student populations, chapters 5 and 6 used students and academic staff members only. This was done for practical reasons as no athlete sample could be immediately recruited to participate in the studies. Thus, the results of the intervention studies might not generalise to elite athlete populations, which is a common issue observed in CAT intervention studies (see chapter 2). Furthermore, the slightly artificial laboratory-based competitive setting may also have detracted from ecological validity. Next to the low generalisability of the findings to athlete populations, the laboratory-based setting might also have provoked less task engagement than a real-world motivated performance situation.

8.4 Directions for Future Research

This thesis could inspire future research in several ways. The field of CAT would benefit from more systematic reviews and/or meta-analyses similar to the one reported in chapter 2. For example, outcomes other than performance (e.g., cognition, emotion, motivation, mental health, & physical health) could be examined in order to test the theoretical predictions of the BPSM and related theoretical models (Blascovich, 2008a; Blascovich, 2008b; Jones, Sheffield, Meijen, & McCarthy, 2009; Vine et al., 2016). As Blascovich (2008b) presented a theoretical pathway from repetitive threat state experience to cardiovascular disease, a systematic review or meta-analysis could for example summarise the state of empirical research in this highly relevant domain (World Health Organisation, 2017). Another review could examine emotional outcomes of CAT states as specified by the TCTSA (Jones et al., 2009). As the TCTSA only specifies emotions of positive (versus negative) valence, and facilitative (versus non-facilitative) interpretations of emotions to be experienced in a challenge (versus a threat) state, future work could also more precisely summarise what emotions have been associated with a challenge (versus a threat) state in the literature. The field could equally benefit from a systematic review and/or meta-analysis of CAT manipulations and challenge-promoting interventions, as a considerable number of studies has been published thus far reporting on the effects of interventions on CAT states (e.g., Feinberg & Aiello; Moore et al., 2015; Sammy et al., 2017; Williams & Cumming, 2012).

The findings showing that CAT states appear to vary largely as a function of personal and person by situation-interactional factors could direct new research toward the identification of personality traits (e.g., hardiness; Kobasa, Maddi, & Kahn, 1982), other psychological (e.g., self-esteem; Seery, Blascovich, Weisbuch, & Vick, 2004), or physiological dispositions (e.g., hormonal; Oliveira & Oliveira, 2014; Salvador & Costa, 2009) as antecedents or correlates of CAT states. In addition, research could investigate what processes could be responsible for the person by situation interaction effects observed. For example, a person by situation interaction could be due to personal skills (i.e., previous experience) interacting with situational aspects like social evaluation (e.g., Blascovich et al., 1999), but there are other possibilities for such interactions worth identifying and exploring.

Regarding the interrelationships between cognitive CAT evaluations, cardiovascular CAT responses, and performance, this thesis has uncovered a need for more research on potential moderators of these relationships that were not previously specified by the BPSM. For example, an observational study could examine age by recruiting a child- and adolescent sample and comparing it with an adult sample regarding the relationships between cognitive CAT and performance. If done in a laboratory setting, the study could involve novel tasks in order to control for the potential confounder of previous experience. In a field setting, experience in the examined performance outcome could be recorded as a control variable and statistically controlled for. Also, the deliberations of Blascovich and Mendes (2000) could be used as a starting point for research into moderators of the relationship between cognitive and cardiovascular indicators of CAT states. Example moderators could include age (see chapter 4) or trait self-reflection and insight (Grant, Franklin, & Langford, 2002), although many more variables should be considered. Testing these moderators could indicate whether the predictions of the BPSM can be supported after taking into account the key moderators, or whether the model would need to be fundamentally revised after a continued lack of association between cognitive and cardiovascular indicators of CAT states.

The experiments reported in this thesis also highlighted a need for closer examination of the relationship between CAT states and performance as a function of interventions. Future studies could test more established interventions (e.g., imagery, Williams, Cumming, & Balanos, 2010; arousal reappraisal, Moore et al., 2015) for potential interactions with CAT states, grouping interactions as either synergistic (i.e., exacerbating performance differences between CAT states), additive (not affecting the relationship between CAT states and performance), or antagonistic (reducing performance differences between CAT states). New intervention studies could also look for actual challenge-promoting interventions, as the interventions tested in this thesis project only influenced the relationship between CAT states and performance, but did not directly act on CAT states by promoting a challenge state. In this context, field research should be prioritised, as all CAT intervention studies reviewed in chapter 2 have been conducted in laboratory settings and all except one (Moore, Wilson, et al., 2013) have used student populations.

8.5 Conclusion

This thesis contributed to the existing literature around the BPSM by examining some previously untested associations and interventions, and by highlighting some important methodological issues surrounding the BPSM. In conclusion, the extant literature suggests that a challenge state is generally superior to a threat state across different performance outcomes and research designs, indicating the relevance of CAT states for sport psychologists and other professionals. Two studies presented in this thesis indicated that CAT states vary largely as a function of personal and person by situation interactional factors, revealing such factors as the most promising targets for potential challenge-promoting interventions. Two further studies showed that two previously supported interventions did not directly promote a challenge state, but rather showed potential in mitigating the negative effects of a threat state, indicating that an awareness of interactions may be helpful in evaluating CAT interventions. Although the systematic review at the beginning of this research project supported the relationships between cognitive and cardiovascular indicators of CAT states and performance as predicted by the BPSM, the relationships were only partly supported by the empirical research presented by this thesis, which provoked the question of whether the predictions of the BPSM require the specification of additional moderators. The thesis highlighted directions for future research as to conducting more systematic reviews and/or meta-analyses of outcomes associated with CAT states; conducting more research into dispositional variables and person by situation interactions; examining how interventions impact CAT states and their relationship with performance (preferably in an applied sports context); and more closely examining and specifying (moderators of) the relationships between cognitive and cardiovascular indicators of CAT states and performance. Taken together, this thesis advanced the CAT literature by providing important evidence relevant to existing questions and pointing out several new questions that had not previously been considered by researchers using the BPSM to study CAT states.

References

Allen, K. M., & Blascovich, J. (1994). Effects of music on cardiovascular reactivity among surgeons. *Journal of American Medical Association, 272*, 882–884. doi: 10.1001/jama.1994.03520110062030

Altman, D. G., & Royston, P. (2006). The cost of dichotomising continuous variables. *BMJ (Clinical Research Ed.), 332*, 1080. doi: 10.1136/bmj.332.7549.1080

Armony, J. L., Corbo, V., Clément, M. H., Brunet, A. (2005). Amygdala response in patients with acute PTSD to masked and unmasked emotional facial expressions. *American Journal of Psychiatry, 162*, 1961–1963. doi: 10.1176/appi.ajp.162.10.1961

Bandura, A., Caprara, G. V., Barbaranelli, C., Gerbino, M., & Pastorelli, C. (2003). Role of affective self-regulatory efficacy in diverse spheres of psychosocial functioning. *Child Development*, *74*, 769–782. doi: 10.1111/1467-8624.00567

Beck, A. T., & Clark, D. A. (1988). Anxiety and depression: An information processing perspective. *Anxiety Research, 1*, 23–36. doi: 10.1080/10615808808248218

Behnke, M., & Kaczmarek, L. D. (2018). Successful performance and cardiovascular markers of challenge and threat: A meta-analysis. *International Journal of Psychophysiology*, *130,* 73–79. doi:10.1016/j.ijpsycho.2018.04.007

Blascovich, J. (2008a). Challenge and threat. In A. J. Elliot (Ed.), *Handbook of approach and avoidance motivation* (pp. 431–445). New York, NY, US: Psychology Press.

Blascovich, J. (2008b). Challenge, threat, and health. In J. Y. Shah, W. L. Gardner (Eds.), *Handbook of motivation science* (pp. 481–493). New York, NY, US: Guilford Press.

Blascovich, J., & Mendes, W. B. (2000). Challenge and threat appraisals: The role of affective cues. In J. P. Forgas (Ed.), *Feeling and thinking: The role of affect in social cognition* (pp. 59–82). New York, NY, US: Cambridge University Press.

Blascovich, J., Mendes, W. B., Hunter, S. B., Lickel, B., & Kowai-Bell, N. (2001). Perceiver threat in social interactions with stigmatized others. *Journal of Personality and Social Psychology, 80,* 253–267. doi: 10.1037/0022-3514.80.2.253

Blascovich, J., Mendes, W. B., Hunter, S. B., & Salomon, K. (1999). Social "facilitation" as challenge and threat. *Journal of Personality and Social Psychology, 77,* 68–77. doi: 10.1037/0022-3514.77.1.68

Blascovich, J., Mendes, W. B., Tomaka, J., Salomon, K., & Seery, M. (2003). The robust nature of the biopsychosocial model challenge and threat: A reply to Wright and Kirby. *Personality and Social Psychology Review, 7,* 234–243. doi: 10.1207/S15327957PSPR0703_03

Blascovich, J., Seery, M. D., Mugridge, C. A., Norris, R. K., & Weisbuch, M. (2004). Predicting athletic performance from cardiovascular indexes of challenge and threat. *Journal of Experimental Social Psychology*, *40*, 683–688. doi: 10.1016/j.jesp.2003.10.007

Blascovich, J., & Tomaka, J. (1996). The biopsychosocial model of arousal regulation. *Advances in experimental social psychology, 28,* 1–51. doi: 10.1016/S0065-2601(08)60235-X

Brennan, R. L. (2011). Generalisability theory and classical test theory. *Applied Measurement in Education, 24*, 1-21. doi: 10.1080/08957347.2011.532417

Brimmell, J., Parker, J., Wilson, M. R., Vine, S. J., & Moore, L. J. (2019). Challenge and threat states, performance, and attentional control during a pressurized soccer penalty task. *Sport, Exercise, and Performance Psychology, 8*, 63–79. doi: 10.1037/spy0000147

Brownley, K. A., Hurwitz, B. E., & Schneiderman, N. (2000). Cardiovascular psychophysiology. In J. T. Cacioppo, L. G. Tassinary, & G. G. Berntson (Eds.), *Handbook of psychophysiology* (pp. 224–264). New York, NY, US: Cambridge University Press.

Burt, K. B., & Obradović, J. (2013). The construct of psychophysiological reactivity: Statistical and psychometric issues. *Developmental Review*, *33*, 29–57. doi:10.1016/j.dr.2012.10.002

Carlsson, A., & Lindqvist, M. (1978). Dependence of 5-HT and catecholamine synthesis on concentrations of precursor amino-acids in rat brain. *Naunyn-Schmiedeberg's Archives of Pharmacology*, *303*, 157–164. doi: 10.1007/BF00508062

Cawley, E. I., Park, S., aan het Rot, M., Sancton, K., Benkelfat, C., Young, S. N., & ... Leyton, M. (2013). Dopamine and light: Dissecting effects on mood and motivational states in women with subsyndromal seasonal affective disorder. *Journal of Psychiatry & Neuroscience*, *38*, 388–397. doi:10.1503/jpn.120181

Chalabaev, A., Major, B., Cury, F., & Sarrazin, P. (2009). Physiological markers of challenge and threat mediate the effects of performance-based goals on performance. *Journal of Experimental Social Psychology*, *45*, 991–994. doi: 10.1016/j.jesp.2009.04.009

Chalabaev, A., Major, B., Sarrazin, P., & Cury, F. (2012). When avoiding failure improves performance: Stereotype threat and the impact of performance goals. *Motivation and Emotion, 36,* 130–142. doi: 10.1007/s11031-011-9241-x

Cohen, J. (1992). A power primer. *Psychological Bulletin*, *112*, 155–159. doi: 10.1037/0033-2909.112.1.155

Colzato, L. S., Jongkees, B. J., Sellaro, R., & Hommel, B. (2013). Working memory reloaded: Tyrosine repletes updating in the N-Back task. *Frontiers in Behavioral Neuroscience, 7*, 1–5. doi: 10.3389/fnbeh.2013.00200

Cools, R. (2006). Dopaminergic modulation of cognitive function-implications for L-DOPA treatment in Parkinson's disease. *Neuroscience and Biobehavioral Reviews, 30*, 1–23. doi: 10.1016/j.neubiorev.2005.03.024

Corbetta, M., & Shulman, G. L. (2002). Control of goal-directed and stimulus-driven attention in the brain. *Nature Reviews Neuroscience, 3,* 201–215. doi: 10.1038/nrn755

Cronbach, L. J., & Furby, L. (1970). How should we measure ‘‘change’’— or should we? *Psychological Bulletin, 74*, 68–80. doi: 10.1037/h0029382

Dienstbier, R. A. (1989). Arousal of physiological toughness: Implications for physical and mental health. *Psychological Review, 96,* 84–100. doi: 10.1037/0033-295X.96.1.84

Drach-Zahavy, A., & Erez, M. (2002). Challenge versus threat effects on the goal-performance relationship. *Organizational Behavior and Human Decision Processes, 88*, 667–682. doi: 10.1016/S0749-5978(02)00004-3

Ell, S. W., Cosley, B., & McCoy, S. K. (2011). When bad stress goes good: increased threat reactivity predicts improved category learning performance. *Psychonomic Bulletin & Review, 18*, 96–102. doi: 10.3758/s13423-010-0018-0

Fédération Internationale de Gymnastique. (2017, November 1). 2017-2020 Code of Points. Retrieved from: http://www.fig-gymnastics.com/site/rules/disciplines/tra

Feinberg, J. M., & Aiello, J. R. (2010). The effect of challenge and threat appraisals under evaluative presence. *Journal of Applied Social Psychology*, *40*, 2071–2104. doi: 10.1111/j.1559-1816.2010.00651.x

Fernstrom, J., & Fernstrom, M. (2007). Tyrosine, phenylalanine, and catecholamine synthesis and function in the brain. *The Journal of Nutrition*, *137*, 1539S.

Freeman, P., & Rees, T. (2009). How does perceived support lead to better performance? An examination of potential mechanisms. *Journal of Applied Sport Psychology*, *21*, 429–441. doi: 10.1080/10413200903222913

Gildea, K. M., Schneider, T. R., & Shebilske, W. L. (2007). Stress appraisals and training performance on a complex laboratory task. *Human Factors*, *49*, 745–758. doi: 10.1518/001872007X215818

Glass, G. V. (1976). Primary, Secondary, and Meta-Analysis of Research. *Educational Researcher, 5,* 3–8. doi: 10.2307/1174772

Goglia, L., Tosi, V., Sanchez, A. M., Flamini, M. I., Fu, X.-D., Zullino, S., … Simoncini, T. (2010). Endothelial regulation of eNOS, PAI-1 and t-PA by testosterone and dihydrotestosterone in vitro and in vivo. *Molecular Human Reproduction*, *16*, 761–769. doi: 10.1093/molehr/gaq049

Grant, A. M., Franklin, J., & Langford, P. (2002). The Self-Reflection and Insight Scale: A new measure of private self-consciousness. *Social Behavior and Personality: An International Journal, 30*, 821–835. doi: 10.2224/sbp.2002.30.8.821

Gross, J., Lakey, B., Lucas, J. L., LaCross, R., Plotkowski, A. R., & Winegard, B. (2015). Forecasting the student–professor matches that result in unusually effective teaching. *British Journal of Educational Psychology, 85*, 19–32. doi: 10.1111/bjep.12049

Hardy, J. (2006). Speaking clearly: A critical review of the self-talk literature. *Psychology of Sport and Exercise, 7*, 81–97. doi: 10.1016/j.psychsport.2005.04.002

Hardy, J., Hall, C. R., Gibbs, C., & Greenslade, C. (2005). Self-talk and gross motor skill performance. *Athletic Insight: The Online Journal of Sport Psychology, 7*.

Hardy, J., Oliver, E., & Tod, D. (2009). A framework for the study and application of self-talk within sport. In S. D. Mellalieu & S. Hanton (Eds.), *Advances in Applied Sport Psychology: A Review* (pp. 37–74). New York, NY: Routledge. doi: 10.4324/9780203887073

Harmer, C. J., McTavish, S. B., Clark, L., Goodwin, G. M., & Cowen, P. J. (2001). Tyrosine depletion attenuates dopamine function in healthy volunteers. *Psychopharmacology, 154*, 105–111. doi:10.1007/s002130000613

Hartley, T. R., Lovallo, W. R., & Whitsett, T. L. (2004). Cardiovascular effects of caffeine in men and women. *The American Journal of Cardiology, 93*, 1022–1026. doi: 10.1016/j.amjcard.2003.12.057

Hase, A., Jung, S. E., & Aan Het Rot, M. (2015). Behavioral and cognitive effects of tyrosine intake in healthy human adults. *Pharmacology, Biochemistry, and Behavior, 133,* 1–6. doi:10.1016/j.pbb.2015.03.008

Hase, A., Gorrie-Stone, T., & Freeman, P. (2019). Tyrosine intake and cardiovascular responses in a motivated performance situation. *Sport, Exercise, and Performance Psychology*, *8,* 80–92. doi: 10.1037/spy0000144.supp

Hase, A., O'Brien, J., Moore, L. J., & Freeman, P. (2019). The relationship between challenge and threat states and performance: A systematic review. *Sport, Exercise, and Performance Psychology*, *8,* 123–144. doi:10.1037/spy0000132

Hatzigeorgiadis, A., Zourbanos, N., Galanis, E., & Theodorakis, Y. (2011). Self-talk and sports performance: A meta-analysis. *Perspectives on Psychological Science*, *6*, 348–356. doi: 10.1177/1745691611413136

Hatzigeorgiadis, A., Zourbanos, N., Latinjak, A. T., & Theodorakis, Y. (2014). Self-talk. In A. G. Papaioannou & D. Hackfort (Eds.), *Routledge companion to sport and exercise psychology: Global perspectives and fundamental concepts.* (pp. 372–385). New York, NY: Routledge/Taylor & Francis Group.

Hatzigeorgiadis, A., Zourbanos, N., Mpoumpaki, S., & Theodorakis, Y. (2009). Mechanisms underlying the self-talk-performance relationship: The effects of motivational self-talk on self-confidence and anxiety. *Psychology of Sport and Exercise*, *10*, 185–192. doi:10.1016/j.psychsport.2008.07.009

Higgins, J. P., & Altman, D. G. (2008). Assessing risk of bias in included studies. In J. P. Higgins & S. Green (Ed.), *Cochrane Handbook for Systematic Reviews of Interventions: Cochrane Book Series*. Retrieved from http://handbook-5-1.cochrane.org.

Hill, D. M., Cheesbrough, M., Gorczynski, P., & Matthews, N. (2019). The Consequences of Choking in Sport: A Constructive or Destructive Experience? *Sport Psychologist, 33*, 12–22. doi: 10.1123/tsp.2018-0070

Hirschl, M. M., Woisetschläger, C., Waldenhofer, U., Herkner, H., & Bur, A. (1999). Finapres vs portapres. *Journal of Human Hypertension, 13*, 899. doi: 10.1038/sj.jhh.1000781

Hunter, S. B. (2001). *Performance under pressure: The impact of challenge and threat states on information processing*. Unpublished doctoral dissertation, University of California, Santa Barbara.

Imholz, B. P. M., Langewouters, G. J., Vanmontfrans, G. A., Parati, G., Vangoudoever, J., Wesseling, K. H., … & Mancia, G. (1993). Feasibility of ambulatory, continuous 24-hour finger arterial pressure recording. *Hypertension, 21,* 65–73. doi: 10.1161/01.HYP.21.1.65

Jamieson, J. P., Crum, A. J., Goyer, J. P., Marotta, M. E., & Akinola, M. (2018). Optimizing stress responses with reappraisal and mindset interventions: an integrated model. *Anxiety, Stress, and Coping, 31,* 245–261. doi:10.1080/10615806.2018.1442615

Jamieson, J. P., Nock, M. K., & Mendes, W. B. (2012). Mind over matter: Reappraising arousal improves cardiovascular and cognitive responses to stress. *Journal of Experimental Psychology: General, 141*, 417–422. doi: 10.1037/a0025719

Jones, M., Meijen, C., McCarthy, P. J., & Sheffield, D. (2009). A theory of challenge and threat states in athletes. *International Review of Sport and Exercise Psychology, 2*, 161–180. doi:10.1080/17509840902829331

Jongkees, B. J., Hommel, B., & Colzato, L. S. (2014). People are different: Tyrosine's modulating effect on cognitive control in healthy humans may depend on individual differences related to dopamine function. *Frontiers in Psychology*, *5*. doi: 10.3389/fpsyg.2014.01101

Kassam, K. S., Koslov, K., & Mendes, W. B. (2009). Decisions under distress: Stress profiles influence anchoring and adjustment. *Psychological Science*, *20*, 1394–1399. doi: 10.1111/j.1467-9280.2009.02455.x

Kelsey, R. M., Blascovich, J., Leitten, C. L., Schneider, T. R., Tomaka, J., & Wiens, S. (2000). Cardiovascular reactivity and adaptation to recurrent psychological stress: The moderating effects of evaluative observation. *Psychophysiology*, *37*, 748–756. doi: 10.1111/1469-8986.3760748

Kelsey, R. M., Soderlund, K., & Arthur, C. M. (2004). Cardiovascular reactivity and adaptation to recurrent psychological stress: Replication and extension. *Psychophysiology*, *41*, 924–934. doi: 10.1111/j.1469-8986.2004.00245.x

Kenny, D. A., West, T. V., Malloy, T. E., & Albright, L. (2006). Componential analysis of interpersonal perception data. *Personality and Social Psychology Review*, *10*, 282–294. doi: 10.1207/s15327957pspr1004_1

Kim, S. Y., Park, J. E., Lee, Y. J., Seo, H., Sheen, S., Hahn, S., & ... Son, H. (2013). Testing a tool for assessing the risk of bias for nonrandomized studies showed moderate reliability and promising validity. *Journal of Clinical Epidemiology*, *66*, 408–414. doi:10.1016/j.jclinepi.2012.09.016

Kirchner, W. K. (1958). Age differences in short-term retention of rapidly changing information. *Journal of Experimental Psychology*, *55*, 352–358. doi: 10.1037/h0043688

Kishore, K., Ray, K., Anand, J. P., Thakur, L., Kumar, S., & Panjwani, U. (2013). Tyrosine ameliorates heat induced delay in event related potential P300 and contingent negative variation. *Brain and Cognition, 83*, 324–329. doi:10.1016/j.bandc.2013.09.005

Kobasa, S. C., Maddi, S. R., & Kahn, S. (1982). Hardiness and health: A prospective study. *Journal of Personality and Social Psychology, 42*, 168–177. doi: 10.1037/0022-3514.42.1.168

Kruger, J., Dunning, D. (1999). Unskilled and unaware of it: How difficulties in recognizing one's own incompetence lead to inflated self-assessments. *Journal of Personality and Social Psychology, 77*, 1121–1134. doi: 10.1037/0022-3514.77.6.1121

Laborde, S., Lautenbach, F., & Allen, M. S. (2015). The contribution of coping-related variables and heart rate variability to visual search performance under pressure. *Physiology & Behavior*, *139*, 532–540. doi: 10.1016/j.physbeh.2014.12.003

Lakey, B. (2016). Understanding the P×S aspect of within-person variation: A variance partitioning approach. *Frontiers in Psychology*, *6*. doi: 10.3389/fpsyg.2015.02004

Lakey, B., Lutz, C. J., & Scoboria, A. (2004). The information used to judge supportiveness depends on whether the judgment reflects the personality of perceivers, the objective characteristics of targets, or their unique relationships. *Journal of Social and Clinical Psychology*, *23*, 817–835. doi: 10.1521/jscp.23.6.817.54806

Laliberté, E. (2011). metacor: Meta-analysis of correlation coefficients. *R package version 1.0-2.* Retrieved from: http://CRAN.R-project.org/package=metacor

Latinjak, A. T., Hatzigeorgiadis, A., Comoutos, N., & Hardy, J. (2019). Speaking clearly 10 years on: The case for an integrative perspective of self-talk in sport. *Sport, Exercise, and Performance Psychology*. doi: 10.1037/spy0000160.supp

Latinjak, A. T., Torregrossa, M., Comoutos, N., Hernando-Gimeno, C., & Ramis, Y. (2019). Goal-directed self-talk used to self-regulate in male basketball competitions. *Journal of Sports Sciences, 37*, 1429–1433. doi: 10.1080/02640414.2018.1561967

Lazarus, R. S. (1999). *Stress and emotion: A new synthesis*. New York, NY, US: Springer Publishing Co.

Lazarus, R. S., & Folkman, S. (1984). *Stress, appraisal, and coping*. New York: Springer.

Levy, A., Nicholls, A., & Polman, R. (2012). Cognitive appraisals in sport: The direct and moderating role of mental toughness. *International Journal of Applied Psychology*, *2*, 71–76. doi: 10.5923/j.ijap.20120204.05

Lucas, T., Weidner, N., & Janisse, J. (2012). Where does work stress come from? A generalizability analysis of stress in police officers. *Psychology & Health, 27*, 1426–1447. doi: 10.1080/08870446.2012.687738

Mahoney, C. R., Castellani, J., Kramer, F. M., Young, A., & Lieberman, H. R. (2007). Tyrosine supplementation mitigates working memory decrements during cold exposure. *Physiology & Behavior*, *92*, 575–582. doi:10.1016/j.physbeh.2007.05.003

McHorney, C. A., Ware, J. E., Rogers, W., Raczek, A. E., & Lu, J. R. (1992). The validity and relative precision of MOS short- and long-form health status scales and Dartmouth COOP Charts: Results from the medical outcomes study. *Medical Care, 30*, 253–265. doi:10.1097/00005650-199205001-00025

McLean, A., Rubinsztein, J. S., Robbins, T. W., & Sahakian, B. J. (2004). The effects of tyrosine depletion in normal healthy volunteers: Implications for unipolar depression. *Psychopharmacology, 171*, 286–297. doi:10.1007/s00213-003-1586-8

Mendes, W. B., Blascovich, J., Hunter, S. B., Lickel, B., & Jost, J. T. (2007). Threatened by the unexpected: physiological responses during social interactions with expectancy-violating partners. *Journal of Personality and Social Psychology, 92*, 698–716. doi: 10.1037/0022-3514.92.4.698

Mendes, W. B., Blascovich, J., Lickel, B., & Hunter, S. (2002). Challenge and threat during social interaction with white and black men. *Personality and Social Psychology Bulletin, 28*, 939–952. doi: 10.1177/01467202028007007

Mendes, W. B., Major, B., McCoy, S., & Blascovich, J. (2008). How attributional ambiguity shapes physiological and emotional responses to social rejection and acceptance. *Journal of Personality and Social Psychology, 94*, 278–291. doi: 10.1037/0022-3514.94.2.278

Moher, D., Liberati, A., Tetzlaff, J., & Altman, D. G. (2009). Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *Plos Medicine, 6*, e1000097. doi: 10.1016/j.jclinepi.2009.06.005

Molenberghs, P., Trautwein, F., Böckler, A., Singer, T., & Kanske, P. (2016). Neural correlates of metacognitive ability and of feeling confident: A large-scale fMRI study. *Social Cognitive and Affective Neuroscience*, 11, 1942–1951. doi: 10.1093/scan/nsw093

Moncada, S., Palmer, R. M., & Higgs, E. A. (1991). Nitric oxide: physiology, pathophysiology, and pharmacology. *Pharmacological Reviews*, *43*, 109–142.

Moore, L. J., Vine, S. J., Freeman, P., & Wilson, M. R. (2013). Quiet eye training promotes challenge appraisals and aids performance under elevated anxiety. *International Journal of Sport and Exercise Psychology*, *11*, 169–183. doi: 10.1080/1612197X.2013.773688

Moore, L. J., Vine, S. J., Wilson, M. R., & Freeman, P. (2012). The effect of challenge and threat states on performance: An examination of potential mechanisms. *Psychophysiology*, *49*, 1417–1425. doi: 10.1111/j.1469-8986.2012.01449.x

Moore, L. J., Vine, S. J., Wilson, M. R., & Freeman, P. (2014). Examining the antecedents of challenge and threat states: the influence of perceived required effort and support availability. *International Journal of Psychophysiology*, *93*, 267–273. doi: 10.1016/j.ijpsycho.2014.05.009

Moore, L. J., Vine, S. J., Wilson, M. R., & Freeman, P. (2015). Reappraising threat: How to optimize performance under pressure. *Journal of Sport & Exercise Psychology*, *37*, 339–343. doi: 10.1123/jsep.2014-0186

Moore, L. J., Wilson, M. R., Vine, S. J., Coussens, A. H., & Freeman, P. (2013). Champ or chump? Challenge and threat states during pressurized competition. *Journal of Sport & Exercise Psychology*, *35*, 551–562. doi: 10.1123/jsep.35.6.551

Moore, L. J., Young, T., Freeman, P., & Sarkar, M. (2018). Adverse life events, cardiovascular responses, and sports performance under pressure. *Scandinavian Journal of Medicine & Science in Sports, 28*, 340–347. doi:10.1111/sms.12928

Muis, K. R., & Edwards, O. (2009). Examining the stability of achievement goal orientation. *Contemporary Educational Psychology, 34*, 265–277. doi: 10.1016/j.cedpsych.2009.06.003

O'Brien, C., Mahoney, C., Tharion, W., Sils, I., & Castellani, J. (2007). Dietary tyrosine benefits cognitive and psychomotor performance during body cooling. *Physiology & Behavior, 90,* 301–307. doi: 10.1016/j.physbeh.2006.09.027

O'Connor, K. M., Arnold, J. A., & Maurizio, A. M. (2010). The prospect of negotiating: Stress, cognitive appraisal, and performance. *Journal of Experimental Social Psychology, 46*, 729–735. doi:10.1016/j.jesp.2010.04.007

Oliveira, G., & Oliveira, R. (2014). Androgen responsiveness to competition in humans: the role of cognitive variables. *Neuroscience and Neuroeconomics, 3,* 19–32. doi: 10.2147/NAN.S55721

Olkin, I., & Pratt, J. W. (1958). Unbiased estimation of certain correlation coefficients. *The Annals of Mathematical Statistics*, *29,* 201–211. Retrieved from https://www.jstor.org/stable/2237306

Peňáz, J. (1973). Photoelectric measurement of blood pressure, volume and flow in the finger. In *Digest of the 10th international conference on medical and biological engineering* (p. 104).

Quigley, K. S., Barrett, L. F., & Weinstein, S. (2002). Cardiovascular patterns associated with threat and challenge appraisals: a within-subjects analysis. *Psychophysiology*, *39,* 292–302. doi: 10.1017/S0048577201393046

Rees, T., Freeman. P., Bell, S., & Bunney, R. (2012). Three generalizability studies of the components of perceived coach support. *Journal of Sport & Exercise Psychology*, *34,* 238–251. doi: 10.1123/jsep.34.2.238

Rith-Najarian, L. R., McLaughlin, K. A., Sheridan, M. A., & Nock, M. K. (2014). The biopsychosocial model of stress in adolescence: Self-awareness of performance versus stress reactivity. *Stress: The International Journal on the Biology of Stress, 17*, 193–203. doi: 10.3109/10253890.2014.891102

Roberts, M. J., Gale, T. E., McGrath, J. S., & Wilson, M. R. (2015). Rising to the challenge: Acute stress appraisals and selection centre performance in applicants to postgraduate specialty training in anaesthesia. *Advances in Health Sciences Education, 21*, 323–339. doi: 10.1007/s10459-015-9629-6

Roiser, J. P., McLean, A., Ogilvie, A. D., Blackwell, A. D., Bamber, D. J., Goodyer, I., & ... Sahakian, B. J. (2005). The subjective and cognitive effects of acute phenylalanine and tyrosine depletion in patients recovered from depression. *Neuropsychopharmacology, 30*, 775–785. doi: 10.1038/sj.npp.1300659

Rosenthal, R. (1979). The file drawer problem and tolerance for null results. *Psychological Bulletin, 86*, 638–641. doi: 10.1037/0033-2909.86.3.638

Sallis, J. F., Prochaska, J. J., & Taylor, W. C. (2000). A review of correlates of physical activity of children and adolescents. *Medicine and Science in Sports and Exercise, 32*, 963–975. doi: 10.1097/00005768-200005000-00014

Salvador, A., & Costa, R. (2009). Coping with competition: Neuroendocrine responses and cognitive variables. *Neuroscience and Biobehavioral Reviews, 33*, 160–170. doi: 10.1016/j.neubiorev.2008.09.005

Sammy, N., Anstiss, P. A., Moore, L. J., Freeman, P., Wilson, M. R., & Vine, S. J. (2017). The effects of arousal reappraisal on stress responses, performance and attention. *Anxiety, Stress, &Coping, 30*, 619–629. doi:10.1080/10615806.2017.1330952

Scheepers, D. (2017). Intergroup status differences as challenge or threat: The role of legitimacy. *Group Processes & Intergroup Relations, 20*, 75–90. doi: 10.1177/1368430215595108

Schneider, T. R. (2004). The role of neuroticism on psychological and physiological stress responses. *Journal of Experimental Social Psychology*, *40*, 795–804. doi: 10.1016/j.jesp.2004.04.005

Schneider, T. R. (2008). Evaluations of stressful transactions: what's in an appraisal? *Stress and Health, 24,* 151–158. doi: 10.1002/smi.1176

Schneider, T. R., Rench, T. A., Lyons, J. B., & Riffle, R. R. (2012). The influence of neuroticism, extraversion and openness on stress responses. *Stress and Health, 28,* 102–110. doi: 10.1002/smi.1409

Scholl, A., Moeller, K., Scheepers, D., Nuerk, H., & Sassenberg, K. (2017). Physiological threat responses predict number processing. *Psychological Research, 81*, 278–288. doi: 10.1007/s00426-015-0719-0

Schulze, R. (2004). *Meta-analysis: A comparison of approaches*. Ashland, OH: Hogrefe & Huber Publishers.

Seery, M. D. (2013). The biopsychosocial model of challenge and threat: Using the heart to measure the mind. *Social and Personality Psychology Compass, 7*, 637–653. doi: 10.1111/spc3.12052

Seery, M. D., Blascovich, J., Weisbuch, M., & Vick, S. B. (2004). The Relationship between self-esteem level, self-esteem stability, and cardiovascular reactions to performance feedback. *Journal of Personality and Social Psychology, 87*, 133–145. doi: 10.1037/0022-3514.87.1.133

Seery, M. D., Weisbuch, M., & Blascovich, J. (2009). Something to gain, something to lose: The cardiovascular consequences of outcome framing. *International Journal of Psychophysiology, 73*, 308–312. doi: 10.1016/j.ijpsycho.2009.05.006

Seery, M. D., Weisbuch, M., Hetenyi, M. A., & Blascovich, J. (2010). Cardiovascular measures independently predict performance in a university course. *Psychophysiology*, *47*, 535–539. doi: 10.1111/j.1469-8986.2009.00945.x

Shavelson, R. J., & Webb, N. M. (2006). Generalizability Theory. In J. L. Green, G. Camilli, & P. B. Elmore (Eds.), *Handbook of complementary methods in education research.* (pp. 309–322). Mahwah, NJ: Lawrence Erlbaum Associates Publishers.

Sherwood, A., Allen, M.T., Kelsey, R.M., Lovallo, W.R., & van Doornen, L.J.P. (1990). Methodological guidelines for impedance cardiography. *Psychophysiology*, *27*, 1–23. doi:10.1111/j.1469-8986.1990.tb02171.x

Shimizu, M., Seery, M. D., Weisbuch, M., & Lupien, S. P. (2011). Trait social anxiety and physiological activation: Cardiovascular threat during social interaction. *Personality and Social Psychology Bulletin*, *37,* 94–106. doi:10.1177/0146167210391674

Shurtleff, D., Thomas, J. R., Schrot, J., Kowalski, K., & Harford, R. (1994). Tyrosine reverses a cold-induced working memory deficit in humans. *Pharmacology, Biochemistry and Behavior, 47,* 935–941. doi:10.1016/0091-3057(94)90299-2

Skinner, N., & Brewer, N. (2004). Adaptive approaches to competition: Challenge appraisals and positive emotion. *Journal of Sport & Exercise Psychology*, *26*, 283–305. doi: 10.1123/jsep.26.2.283

Smith, R. E. (1989). Effects of coping skills training on generalized self-efficacy and locus of control. *Journal of Personality and Social Psychology*, *56*, 228–233. doi: 10.1037/0022-3514.56.2.228

Stevens, J. P. (2009). *Applied multivariate statistics for the social sciences, 5th ed*. New York, NY, US: Routledge/Taylor & Francis Group.

Streamer, L., Seery, M. D., Kondrak, C. L., Lamarche, V. M., & Saltsman, T. L. (2017). Not I, but she: The beneficial effects of self-distancing on challenge/threat cardiovascular responses. *Journal of Experimental Social Psychology*, *70*, 235–241. doi: 10.1016/j.jesp.2016.11.008

Strüder, H., Hollmann, W., Platen, P., Donike, M., Gotzmann, A., & Weber, K. (1998). Influence of paroxetine, branched-chain amino acids and tyrosine on neuroendocrine system responses and fatigue in humans. *Hormone and Metabolic Research = Hormon- und Stoffwechselforschung = Hormones et Métabolisme, 30*, 188–194. doi: 10.1055/s-2007-978864

Suurmond, R., van Rhee, H., & Hak, T. (2017). Introduction, comparison and validation of Meta-Essentials: A free and simple tool for meta-analysis. *Research Synthesis Methods, 8*, 537–553. doi: 10.1002/jrsm.1260.

Theodorakis, Y., Hatzigeorgiadis, A., & Zourbanos, N. (2012). Cognitions: Self-talk and performance. In S. M. Murphy, S. M. Murphy (Eds.) , *The Oxford handbook of sport and performance psychology* (pp. 191–212). New York, NY: Oxford University Press. doi: 10.1093/oxfordhb/9780199731763.013.0010

Theodorakis, Y., Weinberg, R., Natsis, P., Douma, I., & Kazakas, P. (2000). The effects of motivational versus instructional self-talk on improving motor performance. *The Sport Psychologist*, *14*, 253–272. doi: 10.1123/tsp.14.3.253

Thomas, J. R., Lockwood, P. A., Singh, A., & Deuster, P. A. (1999). Tyrosine improves working memory in a multitasking environment. *Pharmacology, Biochemistry and Behavior*, *64*, 495–500. doi:10.1016/S0091-3057(99)00094-5

Tod, D., Hardy, J., & Oliver, E. (2011). Effects of self-talk: A systematic review. *Journal of Sport & Exercise Psychology*, *33*, 666–687. doi: 10.1123/jsep.33.5.666

Tomaka, J., Blascovich, J., Kelsey, R., & Leitten, C. (1993). Subjective, physiological, and behavioral effects of threat and challenge appraisal. *Journal of Personality and Social Psychology*, *65*, 248–260. doi: 10.1037/0022-3514.65.2.248

Tomaka, J., Blascovich, J., Kibler, J., & Ernst, J. M. (1997). Cognitive and physiological antecedents of threat and challenge appraisal. *Journal of Personality and Social Psychology*, *73*, 63–72. doi: 10.1037/0022-3514.73.1.63

Tomaka, J., Palacios, R. L., Champion, C., & Monks, S. (2018). Development and validation of an instrument that assesses individual differences in threat and challenge appraisal. *Journal of Depression and Anxiety*, *7*, 313. doi: 10.4172/2167-1044.1000313

Trotman, G. P., Williams, S. E., Quinton, M. L., & Veldhuijzen van Zanten, J. C. S. (2018). Challenge and threat states: examining cardiovascular, cognitive and affective responses to two distinct laboratory stress tasks. *International Journal of Psychophysiology*, *126*, 42–51. doi: 10.1016/j.ijpsycho.2018.02.004

Tumilty, L., Davison, G., Beckmann, M., & Thatcher, R. (2014). Failure of oral tyrosine supplementation to improve exercise performance in the heat. *Medicine and Science in Sports and Exercise*, *46*, 1417–1425. doi: 10.1249/MSS.0000000000000243

Turner, M. J., Jones, M. V, Sheffield, D., Barker, J. B., & Coffee, P. (2014). Manipulating cardiovascular indices of challenge and threat using resource appraisals. *International Journal of Psychophysiology*, *94*, 9–18. doi: 10.1016/j.ijpsycho.2014.07.004

Turner, M. J., Jones, M. V, Sheffield, D., Slater, M. J., Barker, J. B., & Bell, J. J. (2013). Who thrives under pressure? Predicting the performance of elite academy cricketers using the cardiovascular indicators of challenge and threat states. *Journal of Sport & Exercise Psychology*, *35,* 387–397. doi: 10.1123/jsep.35.4.387

Turner, M. J., Jones, M. V., Sheffield, D., & Cross, S. L. (2012). Cardiovascular indices of challenge and threat states predict competitive performance. *International Journal of Psychophysiology*, *86*, 48–57. doi: 10.1016/j.ijpsycho.2012.08.004

van de Rest, O., Bloemendaal, M., de Heus, R., & Aarts, E. (2017). Dose-dependent effects of oral tyrosine administration on plasma tyrosine levels and cognition in aging. *Nutrients, 9*, 1279. doi: 10.3390/nu9121279

Van Raalte, J. L., Vincent, A., & Brewer, B. W. (2016). Self-talk: Review and sport-specific model. *Psychology of Sport and Exercise*, *22*, 139–148. doi: 10.1016/j.psychsport.2015.08.004

Vine, S. J., Freeman, P., Moore, L. J., Chandra-Ramanan, R., & Wilson, M. R. (2013). 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. *Journal of Experimental Psychology: Applied*, *19*, 185–194. doi: 10.1037/a0034106

Vine, S. J., Moore, L. J., & Wilson, M. R. (2016). An integrative framework of stress, attention, and visuomotor performance. *Frontiers in Psychology, 7.* doi: 10.3389/fpsyg.2016.01671

Vine, S. J., Uiga, L., Lavric, A., Moore, L. J., Tsaneva-Atanasova, K., & Wilson, M. R. (2015). Individual reactions to stress predict performance during a critical aviation incident. *Anxiety, Stress, & Coping*, *28*, 467–477. doi: 10.1080/10615806.2014.986722

Wadey, R., & Hanton, S. (2008). Basic psychological skills usage and competitive anxiety responses: Perceived underlying mechanisms. *Research Quarterly for Exercise and Sport*, *79*, 363–373. doi: 10.1080/02701367.2008.10599500

Wesseling, K. H. (1996). Finger arterial pressure measurement with Finapres. *Zeitschrift Für Kardiologie*, *85*, 38–44. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/8896298

White, J. B. (2008). Fail or flourish? Cognitive appraisal moderates the effect of solo status on performance. *Personality and Social Psychology Bulletin*, *34*, 1171–1184. doi: 10.1177/0146167208318404

Williams, S., & Cumming, J. (2012). Challenge vs. threat: Investigating the effect of using imagery to manipulate stress appraisal of a dart throwing task. *Sport & Exercise Psychology Review*, *8*, 4–21.

Williams, S. E., Cumming, J., & Balanos, G. M. (2010). The use of imagery to manipulate challenge and threat appraisal States in athletes. *Journal of Sport & Exercise Psychology*, *32*, 339–358. doi: 10.1123/jsep.32.3.339

World Health Organization (2017, May 17). *Cardiovascular diseases (CVDs)*. Retrieved from https://www.who.int/en/news-room/fact-sheets/detail/cardiovascular-diseases-(cvds)

Wright, R. A., & Kirby, L. D. (2003). Cardiovascular correlates of challenge and threat appraisals: A critical examination of the biopsychosocial analysis. *Personality* *and Social Psychology Review*, *7*, 216-233. doi: 10.1207/S15327957PSPR0703_02

Zanstra, Y. J., Johnston, D. W., & Rasbash, J. (2010). Appraisal predicts hemodynamic reactivity in a naturalistic stressor. *International Journal of Psychophysiology*, *77*, 35–42. doi: 10.1016/j.ijpsycho.2010.04.00

Appendix A

Table A1

N-Back Task Items by Week (Chapter 3).

Order	Week 1	Week 2	Week 3
1	A	B	E
2	T	U	W
3	D	E	A
4	A	B	G
5	T	U	V
6	Y	Z	A
7	E	F	G
8	D	E	F
9	Z	A	B
10	E	F	G
11	F	G	H
12	Z	A	B
13	A	B	C
14	F	G	H
15	V	W	X
16	A	B	C
17	S	T	U
18	V	W	X
19	T	U	V
20	Z	A	B
21	V	W	X
22	T	U	V
23	Z	A	B

Appendix B

Table B1

Subtraction Task Items by Week (Chapter 3)

Week 1					Week 2					Week 3				
Exercise	Answer Options				Exercise	Answer Options				Exercise	Answer Options			
	C	D1	D2	D3		C	D1	D2	D3		C	D1	D2	D3
412 – 122	290	298	390	324	522 – 412	110	134	10	94	822 – 712	110	134	10	94
524 – 371	153	143	163	165	635 – 242	393	413	377	417	535 – 342	193	213	277	117
174 – 121	53	55	153	155	286 – 232	54	58	48	46	488 – 434	54	58	48	46
899 – 672	227	231	131	271	699 – 572	127	227	172	272	599 – 472	127	227	172	272
915 – 328	587	685	683	583	825 – 328	497	503	597	407	525 – 228	297	303	397	207
537 – 497	40	30	44	34	727 – 297	430	574	474	470	627 – 197	430	574	474	470
126 – 112	14	16	18	22	156 – 108	48	96	38	148	186 – 138	48	96	38	148
892 – 624	268	277	267	278	902 – 424	478	377	367	378	912 – 434	478	377	367	378
143 – 112	31	32	35	29	173 – 152	21	22	25	29	293 – 272	21	22	25	29
475 – 219	256	266	254	264	675 – 319	356	366	354	364	563 – 319	244	122	366	488
429 – 357	72	67	66	76	229 – 157	72	67	66	76	559 – 157	402	416	316	392
926 – 921	15	17	21	27	736 – 721	15	17	21	27	436 – 321	115	117	121	127
524 – 426	98	89	102	92	648 – 214	434	442	432	424	648 – 414	234	242	232	224
744 – 511	233	255	235	253	534 – 511	23	25	13	15	434 – 411	23	25	13	15
745 – 289	456	536	436	476	845 – 269	576	536	436	476	855 – 279	576	536	436	476
670 – 105	565	575	535	475	240 – 105	135	45	145	35	480 – 105	375	345	275	385
508 – 428	80	96	106	90	608 – 428	180	136	236	124	808 – 628	180	136	236	124
674 – 659	15	5	13	23	784 – 769	15	75	13	23	384 – 369	15	73	13	23
420 – 301	119	89	121	181	820 – 501	319	289	321	281	725 – 501	224	226	221	225
783 – 763	20	23	26	30	983 – 943	40	43	46	49	693 – 633	60	63	66	69

Note. C = Correct answer option. D1 = Distractor option 1. D2 = Distractor option 2. D3 = Distractor option 3.

Appendix C

Table C1

N-Back Task Items by Week (Chapter 6).

Order	Week 1	Week 2
1	F	G
2	X	Y
3	B	C
4	H	I
5	W	X
6	B	C
7	H	I
8	G	H
9	C	D
10	H	I
11	I	J
12	C	D
13	D	E
14	I	J
15	Y	Z
16	D	E
17	V	W
18	Y	Z
19	W	X
20	C	D
21	Y	Z
22	W	X
23	C	D