Development and Validation of the Characteristics of Resilience in Sports Teams Inventory

Steven Decroos[1], Robin L. J. Lines[2], Paul B. C. Morgan[3], David Fletcher[4],

Mustafa Sarkar[5], Katrien Fransen[1], Filip Boen[1], Gert Vande Broek[1]

[1]Physical Activity, Sports and Health Research Group, University of Leuven, Belgium

[2]School of Physiotherapy and Exercise Science, Curtin University, Australia

[3]Centre for Human Performance, Exercise, and Wellbeing, Buckinghamshire New University, United Kingdom

[4]School of Sport, Exercise and Health Sciences, Loughborough University, United Kingdom

[5]School of Science and Technology, Nottingham Trent University, United Kingdom

Author note

This research was funded by the Flemish Government (Department of Culture, Youth, Sport, and Media) as part of the “M-Factor” project, i.e. an experimental project on motivational coaching in youth sports.

Correspondence concerning this article should be addressed to Steven Decroos, Department of Kinesiology, Tervuursevest 101, B-3001 Heverlee, Belgium.

E-mail: steven.decroos@kuleuven.be

Abstract

This multi-study paper reports the development and initial validation of an inventory for the Characteristics of Resilience in Sports Teams (CREST). In four related studies, 1225 athletes from Belgium and the United Kingdom were sampled. The first study provided content validity for an initial item set. The second study explored the factor structure of the CREST, yielding initial evidence but no conclusive results. In contrast, the third and fourth study provided evidence for a two-factor measure, reflecting (a) the team's ability to display resilient characteristics and (b) the vulnerabilities being displayed under pressure. Overall, the CREST was shown to be reliable at the between-players and the between-teams level, as well as over time. Moreover, its concurrent validity was verified by linking the characteristics of team resilience with various relevant team processes. Its discriminant validity was established by comparing the CREST measures with individual athletes' resilient traits. In conclusion, the CREST was argued to be a usable state-like measure of team-level resilient characteristics and vulnerabilities. To gain further understanding of team resilience as a process, this measurement could be used in future process-oriented research examining adverse events and sports team's pre- and post-adversity functioning.

Keywords: team dynamics, pressure, protective factors, stress, questionnaire

Development and Validation of the Characteristics of Resilience in Sports Teams Inventory

Any engagement in competitive sports inevitably coincides with experiences of pressure and adversities such as performance slumps, day-to-day hassles, and hampering life events (Fletcher, Hanton, & Mellalieu, 2006; Mellalieu, Hanton, & Fletcher, 2006). It has even been argued that adverse experiences are not only inevitable features of competitive sport, but that they also constitute a prerequisite to acquire high levels of performance (Collins & MacNamara, 2012; Sarkar, Fletcher, & Brown, 2015). Nonetheless, all pressure and adversities that athletes experience carry the potential to impair their development (Fraser-Thomas & Cote, 2009; Theokas, 2009). For instance, athletes under pressure are more likely to display damaged self-esteem (Gagne, Ryan, & Bargmann, 2003), unsportsmanlike behaviours (Vansteenkiste, Mouratidis, & Lens, 2010), and burnout (Tabei, Fletcher, & Goodger, 2012). Therefore, a desirable challenge for competitive athletes to be successful is to positively adapt to the adversities they encounter (Fletcher & Sarkar, 2012; Gould, Dieffenbach, & Moffett, 2002). This suggests that athletes' resilience is a key determinant of sporting success.

At the individual level, psychological resilience has been defined as "the role of mental processes and behaviour in promoting personal assets and protecting an individual from the potential negative effect of stressors" (Fletcher & Sarkar, 2012, p. 675; 2013, p. 16). From this process-oriented perspective (cf. Galli & Vealy, 2008), both personal (e.g., self-determining motives and confidence in one's own abilities) and environmental factors (e.g., perceived social support) have been found to generate a variety of resilient outcomes. These outcomes include improved learning and a broadened life perspective (Sarkar & Fletcher, 2014a; 2014b). For an overview of individual athletes' resilience, see Galli and Gonzalez (2015) and Sarkar and Fletcher (2014b).

In teams, however, the study of group-level resilience additionally requires a socio-ecological perspective, which includes the shared experiences in the team environment and the interactive resources that teams can employ (Galli, 2016; Yukelson & Weinberg, 2016). To illustrate, collectively encountered stressors in sports teams include joint logistical problems, letting others down, key player lay-offs compelling strategic adaptations, and interpersonal tensions (cf. Holt & Hogg, 2002; Levy, Polman, Nicholls, & Marchant, 2009). These collective stressors clearly contrast with individual athletes' issues with finance, diets, or their sport-life interface. Nonetheless, it is only recently that scholars have revealed how a team's collective resources can be employed to withstand such shared demands (for a review, see Galli, 2016; Yukelson & Weinberg, 2016).

In the first study of team resilience, Morgan, Fletcher, and Sarkar (2013) defined team resilience as a "dynamic, psychosocial process that protects a group of individuals from the potential negative effects of stressors they collectively encounter" (p. 552). Four main characteristics of resilient sports teams emerged from Morgan et al.'s (2013) study: (1) *group structure* (e.g., reflecting on a shared vision during stressors, open and honest communication as a norm, shared leadership roles); (2) *mastery approaches* towards adversities (e.g., focus on learning and improvement as a group during setbacks, thorough preparation to withstand stressors, and gaining experience of challenging situations); (3) *social capital* (e.g., deep emotional bonds between team members, perceived social support, and the absence of a blame culture when experiencing failures); and (4) *collective efficacy* (e.g., gaining belief from successful past experiences of adversity, sticking together during setbacks, and gaining belief from the acts of team members during stressors).

In addition to these main characteristics that define the resilient state of sports teams, a follow-up study by Morgan, Fletcher, and Sarkar (2015) revealed developmental antecedents of team resilience. In particular, it was revealed that coaches could prepare their teams for upcoming adversities in four ways: (1) by developing a collective vision for functioning under pressure; (2) by using shared experiences to learn from; (3) by empowering a strong team identity; (4) and by promoting positive emotions and enjoyment. Overall, the findings of Morgan et al. (2013, 2015) advanced our knowledge of team resilience by describing what a resilient team looks like (i.e., resilient characteristics) and how a team can grow more resilient over time (i.e., developmental antecedents).

Concerning measurements of resilience, most resilience studies at both the individual and the team level have employed a qualitative design. One exception at the individual level is the research by Gucciardi, Jackson, Coulter, and Mallett (2011), in which the Connor-Davidson Resilience Scale (CD-RISC; Connor & Davidson, 2003) has been used to quantify cricketers' resilience. In this study, the authors (Gucciardi et al., 2011) found support for the validity and reliability of the 10-item CD-RISC in the sports context and highlighted the usefulness of this scale to examine how individual athletes respond positively to adversity. Also, Gonzalez, Moore, Newton, and Galli (2016) recently reported the CD-RISC as a potential tool to measure resilience as a trait characteristic of individual athletes. However, at the team level, no measurement instrument has been developed yet. To our knowledge, no quantitative studies in team sports reported team resilience as a solid and sound measure (cf. Galli, 2016).

Alternatively, one could transpose the recommendation of Sarkar and Fletcher (2013) for individual performers into the team level and use a collection of scales that are indicative of team resilience. For instance, the persistence and effort subscales of the Collective Efficacy Questionnaire for Sports (CEQS; Short, Sullivan, & Feltz, 2005) quantify, respectively, the extent to which athletes believe they will be able to perform under pressure and whether athletes believe in their ability to show a strong work ethic. Similarly, the Observational Collective Efficacy Scale for Sports (OCESS; Fransen, Kleinert, Dithurbide, Vanbeselaere, & Boen, 2014) assesses the processes that enable teams to sustain their levels of team confidence in future performance situations. Also, the Peer Motivational Climate in Youth Sports (PMCYS; Ntoumanis & Vazou, 2005) could provide insights into a resilient team's mastery approach. More specifically, the extent to which athletes value effort, team improvement, and social bonds, or, inversely, the extent to which intra-team conflicts are present, have been argued to resonate with team resilience (Morgan et al., 2013, 2015).

Nonetheless, caution is warranted when constructs that resemble each other nomologically are regarded as structural or functional equivalents (Morgeson & Hofmann, 1999). For example, resilient teams' mastery approaches in dealing with setbacks were labelled as such because of these teams' emphases on improving their behavioural responses setback after setback (Morgan et al., 2013). However, this characteristic also comprises the effectiveness of behavioural responses to setbacks and the teams' flexibility in managing changes. Yet, to our knowledge, these latter components of a teams' mastery approach towards setbacks have not structurally nor functionally been demonstrated within the mastery versus performance framework of the PMCYS. Moreover, the items of the PMCYS, similar to the items of the effort subscale of the CEQS, disregarded stressors and adversities. These items could equally be assessed when thriving or experiencing collective competence. In contrast, the persistence subscale of the CEQS specifically relates to functioning under pressure and, as such, could be hypothesized to converge with team resilience characteristics. Still, the latter persistence subscale lacks integral aspects on handling collective adversities, for example, referring to a team's shared vision.

Therefore and in accordance with D. Chan's (1998) and Morgeson and Hofmann's (1999) guidelines for multilevel research, Morgan et al. (2015) proposed that, "For team resilience research and measurement in sport […] team resilience should be operationalized and assessed differently at different levels of analysis" (p. 99). More specifically, because team resilience not only consists of shared beliefs and group norms, but also comprises behavioural adaptations to stressors and an interactive mastery approach, the absence of agreement on individual athletes' resilient traits does not reflect the absence of team resilience. Moreover, the CD-RISC has already been criticized as a measure for resilience because it excludes interactive and dynamic state-like characteristics (cf. Galli & Gonzalez, 2015).

Instead, a referent-consensus model (cf. D. Chan, 1998, p. 238) is required to reflect the resilient characteristics of a sports team. Such a model uses athletes as informants of their team.

More specifically, a referent-consensus model calculates the athletes' average perception as a "point measurement" of group-level characteristics; and their within-team agreement as the emergent state of those characteristics (Kozlowski, 2015; Lüdtke, Robitzsch, Trautwein, & Kunter, 2009; Marks, Mathieu, & Zaccaro, 2001). To build such a measurement model, the study of Morgan et al. (2013) suits as a starting point. Both the higher-order themes and the quotes of individual athletes constitute empirical inclusion criteria to deduce an initial item set (Johnson, Rosen, & Chang, 2011; Johnson, Rosen, Chang, Djurdjevic, & Taing, 2012).

In addition, such a state-like operationalization of the resilient characteristics of sports teams does not contradict the process-oriented definition of resilience as an umbrella construct (cf. Bonanno, Romero, & Klein, 2015). On the contrary, this referent-consensus model could ultimately contribute to the understanding of team resilience as a process. For example, more process-oriented research examining adverse events and sports teams' pre- and post-adversity functioning could employ multiple state measurements of the resilient characteristics to investigate the impact of within-team variability, convergence over time, and growth trajectories, on team outcomes (cf. Kozlowski, 2015). Moreover, athletes' interactions that reinforce their team's resilience could be reflected throughout the items (cf. Morgeson & Hofmann, 1999). As such, team resilience as an umbrella construct could be partially disentangled by measuring the predictors of resilient outcomes (Bonanno et al., 2015).

In conclusion, to properly and distinctly assess team resilience characteristics in the sports context and, in turn, to advance the knowledge of team functioning during stressors, a specific measure is needed (Galli, 2016; Morgan et al., 2015). Therefore, the purpose of the studies presented in this paper is to develop an inventory for the characteristics of team resilience and to begin the process of validation. The intended inventory is referred to as the Characteristics of Resilience in Sports Teams Inventory (CREST).

Overall Method

As the development and validation of an inventory requires a systematic process (MacKenzie, Podsakoff, & Podsakoff, 2001), four consecutive studies were conducted. Institutional ethical approval was granted by the Loughborough University for Study 1 and the University of Leuven for the consecutive studies. All participants were given an informed consent form before data-acquisition. In Study 4, parents were asked for a signed consent form before approaching the athletes under 16 years old.

Study 1, conducted in the United Kingdom (UK), specifically sought to adhere to the recommendations by Gehlbach and Brinkworth (2011) in scale design. Instead of relying solely on psychometric analysis, Study 1 utilized a variety of developmental techniques, such as multiple expert panels and a multi-language approach. Furthermore, by collaborating with coaches and athletes via expert panels (cf. Dunn, Bouffard, & Rogers, 1999) and cognitive interviews (cf. Dietrich & Ehrlenspiel, 2010), the content validity of the items was optimized before taken further steps.

Consequently, because our referent-consensus model uses single athletes as informants of their teams, we continued the validation process first at the level of measurement. In Studies 2 and 3, single athletes were asked to rate their teams based on the validated set of items from Study 1. The aim of Study 2 was to initially explore the factor structure in a Flemish sample and to test the concurrent validity with specific aspects of the motivational climate (i.e., effort, improvement and relatedness support) and collective efficacy beliefs (i.e., effort and persistence). In Study 3, the factor structure was further explored in a UK sample. Finally, the intended referent-consensus model was administered to full sports teams. More specifically, Study 4 in Flanders aimed at confirming the factorial and construct validity of the CREST at the team level and at testing the intra-team congruence among the coaches' and athletes' perceptions as an emergent state of the team-level resilient characteristics. In the latter two studies (Studies 3 and 4), the discriminant validity of the

CREST was tested against individual athletes' resilient traits measured by the CD-RISC. The concurrent validity was tested with the OCESS and the intra-team conflicts subscale of the PMCYS.

The factor structure of the CREST was explored with Mplus 7.4 (Muthén & Muthén, 1998-2015), following the approach of Morin, Arens, and Marsh (2015). As such, multiple structural equation models were contrasted. First, an independent cluster model in a confirmatory factor analysis (ICM-CFA) was employed to test the structural independence of the four main characteristics comprising the CREST items. Second, possible cross-loadings were evaluated with an exploratory structural equation model (ESEM), based on an orthogonal rotation targeting the four main characteristics. Then, when few cross-loadings would emerge, this ESEM would be further contrasted with a hierarchical ESEM (H-ESEM) to evaluate team resilience as a superordinate factor determining the main characteristics. In contrast, with multiple significant cross-loadings present, a bifactorial ESEM (B-ESEM) would be considered to assess the possibility of a general team resilience factor (G-factor) explaining item level variances apart from the four main characteristics. Finally, an ICM-CFA of the concurrent and/or discriminant measures was added to the resulting H-ESEM or B-ESEM to evaluate the construct validity of the CREST factors. For more details on these procedures, see Marsh, Morin, Parker, and Kaur (2014) and Morin et al. (2015). In study 4, multilevel CFA's were performed to confirm the factor structure of the CREST at both the individual and team level (Brown, 2015).

We opted for a robust full information maximum likelihood estimation because of the categorical Likert nature and non-normal distribution of our data (Muthén & Muthén, 1998-2015). The ESEM solutions were calculated using a targeted orthogonal rotation (Asparouhov & Muthén, 2009). All loadings and cross-loadings were evaluated following the guidelines of Comrey and Lee (.32 = poor, .45 = fair, .55 = good, .63 = very good; 1992). Cut-off criteria for model fit were based on the recommendations of Markland (2007) and Hu and Bentler (1999); more specifically, model fit was accepted if *CFI* and *TLI* attained at least .90 in combination with *RMSEA* ≤ .08 and *SRMR* ≤ .06.

We further hypothesized, conforming to the guidelines of Raykov (2011), that the factor scores of the CREST would correlate no more than .30 with the discriminant measure (i.e., CD-RISC); and that the correlations between the CREST and the concurrent measures would be higher ($r > .30$), but also to remain below the threshold for convergent measures such as the persistence subscale of the CEQS (i.e., $r = .80$). Also in Study 4, additional multilevel mixed models were tested to evaluate whether the team resilience measures merely represented a group of resilient individuals or were yielded by other between-team differences.

Study 1

An initial set of 87 items was based on the study by Morgan et al. (2013). The items were generated with attention to both the higher-order and lower-order themes behind each of the four main characteristics as well as the quotes of athletes. At this stage, a common stem was generated to precede all items with a time reference allowing for a state measure of the resilient characteristics. The stem was “In the past month when my team was under pressure …”. One month equals the reference time of the Organizational Stressor Indicator for Sport Performers (Arnold, Fletcher, & Daniels, 2013). Therefore, this timing was deemed properly long for stressors to have occurred as well as properly short for respondents to remember them. The purpose of this first study was to refine this pool of items and to provide support for their content validity.

Method

Participants and procedures. Three consecutive research actions independently incorporated the opinions of experts in the field and elite team sport athletes to improve the face and content validity of the initial items. First, and in line with other research developing psychological questionnaires (e.g., Arnold et al., 2013; Bolter & Weiss, 2013; Lee, Whitehead, & Ntoumanis, 2007), the initial list of items was submitted to an independent panel of experts. This step was conducted to obtain information on each item’s clarity, relevance, and specificity (Dunn et al., 1999; Gehlbach & Brinkworth, 2011). Forty-six experts from the UK were contacted and fifteen agreed to participate. Eleven were academic experts, had between 5 and 36 years of experience working in academia, and had authored between 7 and 150 publications in international peer reviewed journals. Eight of the eleven academic experts also had between 6 and 34 years of experience in providing sports psychology support in a variety of sports. The other three academic experts were sampled because of their experience in psychometrics. In addition, four experts who primarily worked as applied sport psychologists participated. They had between 13 and 30 years of experience in the field, but they also worked within academia. Three of the four practitioners held a PhD and had between 10 and 50 publications in international peer reviewed journals.

Subsequently, cognitive interviews were conducted to gain insight into how athletes understood, processed, and responded to the generated items (Dietrich & Ehrlenspiel, 2010). Each item was analysed following Conrad and Blair's (2004) classification system for cognitive interviews. The six participating athletes competed in a variety of team sports (i.e., cricket, Gaelic football, and soccer) and were between 22 and 32 years old.

Finally, an expert panel including team sports coaches in addition to sport psychologists, further addressed the findings of the cognitive interviews in two different ways. First, all items were regrouped to reflect which items were closely related in terms of conceptual and practical interpretation. Second, items were evaluated simultaneously in a Flemish and English version; using the translation-back-translation procedure as advocated by Duda and Hayashi (1998).

Measures. The experts were first presented with Morgan et al.'s (2013) definition of team resilience. Then, they were asked to indicate for each item whether they believed it was 'relevant' (i.e., does the item potentially relate to the characteristics of a resilient sports team?), 'clear' (i.e., is this item easily understood?), and 'specific' (i.e., is the item general enough to capture all characteristics in this area?). A part from the tick boxes "yes", "no", and "unsure", the experts were given space to comment on items, to explain their answer, or to make suggestions. This format for collecting expert panel responses has been advocated in previous sport psychology research (e.g., Jowett & Ntoumanis, 2004; Rhind & Jowett, 2010).

In the cognitive interviews, the pool of items was split into two random sets to avoid participant fatigue. Each set reflected a similar number of items representing the four main characteristics of team resilience. The techniques of "think-aloud" and "verbal probing" were used together, as advocated by Willis (2005). With the think-aloud technique, participants were encouraged to think out loud as they answered the questions. In comparison, through verbal probing, interviewees were asked specific questions to gather additional information on items (Willis, 2005). Verbal probes were generally aimed at comprehension/interpretation (e.g., what does the term 'guiding behavioural principles' mean to you?) and at judgment/decision making (e.g., how did you arrive at your answer?).

Results and Discussion

Experts' assessments resulted in 28 items to be retained in their current form, because more than 75% of the experts agreed on the relevance, clarity, and specificity. This threshold is consistent with previous psychometric research employing expert ratings (e.g., Jowett & Ntoumanis, 2004; Rhind & Jowett, 2010). In addition, 9 items that fell beneath the 75% threshold by only one vote were also retained in their current form. These 9 items were retained for subsequent testing, because they lacked additional information of the experts to improve them and, whilst they merely lacked 2% to reach the 75% threshold, they certainly outreached the 50% threshold used in more recent research (e.g., Cronin & Allen, 2017; Pope & Hall, 2014).

For 38 items, the experts' assessments were also close to the 75% threshold but comprised concrete concerns to be retained in their current form. Based on expert recommendations and concerns, these items were revised. For example, 'the team dug in as a group' was reworded to 'the team collectively worked harder'. This adaptation reflected comments from experts, such as; "Dug in would seem quite a culturally driven term and might not translate the same in other English language countries/speakers", and "I don't think that everyone will understand this reference. Perhaps 'worked harder', or 'increased their effort'." Finally, 12 items in which there was an unacceptably low level of agreement (i.e., below 50%) were removed.

The cognitive interviews ranged in duration from 26 to 38 minutes and resulted in the revision of 4 items, while 7 other items were deleted. For example, the item 'team members trusted and respected each other' was split into two separate items reflecting trust and respect. This was done based on comments such as "Trust for me refers to a performance aspect, like trusting your right back to do his job. However, respect is different, it's more performance and outside of it as well, it's almost two questions in one" or "I focus more on trusted. I think respect is different. You can trust someone on the pitch as a player but not respect them". The 7 deleted items were perceived by the athletes as being too ambiguous.

Finally, an expert panel simultaneously evaluated overlapping meanings from the cognitive interviews and discrepancies from the back-translation of Flemish items. For example, the items "thorough preparation helped the team to cope with pressures", "our preparation helped the team to dig deep when under pressure", and "the team was able to adapt to pressures due to proper preparation" were combined in one item, i.e. "thorough preparation helped the team to deal with pressures". In total, 38 items were retained. Each main characteristic was represented by 8 to 10 items (see Table 3 for an overview), complying with the recommendation of (Marsh, 2007) to minimally include 5 items per possible subscale.

Study 2

The primary purpose of this study was to explore the factor structure of the 38-item version of the CREST and to check the conceptual differences between items. The secondary purpose was to test the concurrent validity of the CREST with the effort and persistence subscales of the CEQS and the peer mastery climate subscales (i.e., effort, improvement, and relatedness support) of the PMCYS.

Method

Participants and procedures. Approximately 1800 team sport athletes (from a previous study of some of the authors; Fransen, Vanbeselaere, De Cuyper, Vande Broek, & Boen, 2014), who had previously indicated their benevolence to participate in future research, were contacted via direct mailing. Apart from a possible augmented interest in sport psychology, self-selection bias was not assumed because the previous study (Fransen, Vanbeselaere, et al., 2014) examined a possible classification system for different types of athlete leaders and had no direct links with pressurized situations, setbacks, or resilience. Additional participants were recruited through social media. The inventory was administered online and preceded by an informed consent page.

In total, 711 individuals visited the survey website of which 636 consented to participate. After excluding 8 athletes below the age of 16 and 315 athletes who did not fill out the CREST, a sample of 388 athletes was obtained for further analysis. Participants were all team sport athletes, mainly active in basketball (n = 125) and volleyball (n = 160). Additional demographic information is shown in Table 1.

Measures. The 38-item version of the CREST was administered on two separate webpages containing 19 items each. The items were randomly sorted and the same stem was utilized; "In the past month, when my team was under pressure…" Although the experts in Study 1 suggested using a 6-point scale, we opted for a 7-point scale, still ranging from "*strongly disagree"* to "*strongly agree*". A 7-point scale was preferred over a 6-point scale, because without a mid-point respondents may just conform to their answer on the preceding item (J. C. Chan, 1991; Kampen, 2006; Tourangeau, Couper, & Conrad, 2004). For example, in a study by Garland (1991) with a 5-point scale, only 14% of the participants who previously chose the "*neither/nor*" position, would choose a negative scale point when the mid-point was removed. Also, in psychology research, Malhotra, Krosnick, and Thomas (2009) concluded that "respondents who placed themselves at the midpoint belonged there" (p. 318) because branching the midpoint into directional alternatives yielded no significant gains in criterion validity. As such, omitting a middle alternative would result in respondents being forced to select a response alternative that might not reflect their true attitudinal position and, as a result, reduce the ecological validity.

In addition to the CREST, other scales were administered to assess its concurrent validity. Regarding the CEQS (Short et al., 2005), the subscales for effort (e.g., rate your team's confidence that your team has the ability to demonstrate a strong work ethic) and persistence (e.g., rate your team's confidence that your team has the ability to perform under pressure) were assessed with 4 items each on a 7-point scale ranging from 1 (*strongly disagree*) to 7 (*strongly agree*). In terms of the PMCYS (Ntoumanis & Vazou, 2005), the 5-items subscale for effort (e.g., on this team, most athletes encourage their teammates to try their hardest), the 4-item subscale for improvement (e.g., on this team, most athletes work together to improve the skills they don't do well), and the 3-items subscale for relatedness support (e.g., on this team, most athletes make their teammates feel valued) were also assessed on a 7-point scale (1 = *strongly disagree* to 7 = *strongly agree*). After data acquisition, an ICM CFA revealed a good factorial structure of these concurrent concepts (χ^2 = 327.85, *df* = 166, p < .001, CFI = .95, *TLI* = .94, *RMSEA* = .051, and *SRMR* = .046).

Results and Discussion

The goodness-of-fit indices and the ABIC criteria calculated for the different models are presented in Table 2. Whereas the ICM CFA solution (CFI = .89, TLI = .88) did not reach acceptable fit indices, the ESEM provided a better (ΔSRMR = -.018, ΔABIC = -484) and acceptable (CFI = .92, TLI = .90) fit to the data. Although the indicators in the ICM appeared with high factor loadings (λ > .55, *p* < .001), the ESEM clearly reflected the ICM's bad fit indices with multiple cross-loadings among all indicators. The indicators that were targeted to a mastery approach and the social capital revealed factor loadings between λ = .41 (*p* < .001) and λ = .72 (*p* < .001), mean λ = .58 (*SD* = .10), whereas cross-loadings ranged from λ = .14 (*p* = .64) to λ = .72 (*p* < .001), mean λ = .39 (*SD* = .13). The indicators pertaining to group structure and collective efficacy were mostly insignificant (mean *p* = .58, *SD* = .29) and varied with loadings between λ = .01 (*p* = .99) and λ = .50 (*p* = .49), mean λ = .27 (*SD* = .15); cross-loadings ranged from λ = .01 (*p* = .98) to λ = .41 (*p* = .45), mean λ = .17 (*SD* = .11).

Because the initial four main characteristics did not structurally emerge through the ESEM, a B-ESEM was assessed to evaluate the extent to which team resilience as a G-factor would explain item variance. This B-ESEM additionally improved fit indices (χ^2 = 899.26, *df* = 524, *p* < .001, *CFI* = .95, *TLI* = .93, *RMSEA* = .044, and *SRMR* = .027) and revealed a single factor structure. The loadings of the G-factor ranged from λ = .50 (*p* < .001) and λ = .80 (*p* = < .001), with only 3 indicators below the threshold of λ = .55. In contrast, the loadings of the four main characteristics ranged from λ = -.28 (*p* < .01) and λ = .53 (*p* < .01). The standardized factor loadings of the B-ESEM are presented in Table 3.

Two arguments might support this finding. First, the four main characteristics of team resilience coincided in practice. For example, a sport team's mastery approach towards adversities could not exist without that team's agreed vision on dealing with setbacks or without strong bonds and trust between the teammates. Second, similarity in wording could have produced a large amount of common method variance (cf. Johnson et al., 2011). Because all items were positively worded, 'lazy' respondents could simply have been conforming to their previous answers without rethinking the new situation presented. This problem could occur in questionnaire research which is inherently intrusive and time-consuming (Kozlowski, 2015). A solution to determine the alertness of respondents is to inverse code some of the items, notwithstanding the fact that negative wording in questionnaires has also been criticized (Sliter & Zickar, 2014; Tomas, Oliver, Hontangas, Sancho, & Galiana, 2015). At least, a possible method effect should be addressed in the results and discussion section (Marsh, Scalas, & Nagengast, 2010). Moreover, caution is advised when reverse coding and the use of antonyms is to be preferred over negations. As such, measurement models that demonstrate a sound factorial structure separating items that reflect the presence of a characteristic (e.g., autonomy support) from items reflecting the opposite (e.g., psychological control, which structurally differs from not being autonomy-supportive) can be found also in sport psychology research (e.g., Cruickshank & Collins, 2014).

For the concurrent measures, the ESEM-within-CFA model revealed an acceptable fit to the data (χ^2 = 2264.74, *df* = 1432, *p* < .001, *CFI* = .93, *TLI* = .92, *RMSEA* = .039, and *SRMR* = .037). Also, significant correlations between the concurrent scales and our CREST measurement were revealed. More specifically, the overarching G-factor correlated significantly with the collective belief to put effort in an upcoming game (*r* = .79, *p* < .001) and the perceived peer mastery climate in the team (*r* = .69, *p* < .001). As expected the correlation of the G-factor with the persistence subscale of the CEQS was higher (*r* = .87, *p* < .001) because these items specifically referred to stressors and adversities.

In conclusion, Study 2 provided initial evidence for the CREST as a measure of resilient characteristics in sports teams. One single factor emerged as the most suitable representation of the resilient characteristics. This single factor concurred with the athletes' collective beliefs in their team's ability to demonstrate effort and with a peer mastery climate under normal (non-pressurized) circumstances. It also converged with the athletes' belief in their team's ability to persist when setbacks would occur in the near future. It should be noted though that the single factor solution could have been induced by common method variance residing in the positive wording of all items. Therefore, at this point, a consecutive study was required to further explore the factor structure with some of the items reverse-coded.

Study 3

From the previous results, an optimized 20-item version of the CREST could be obtained by retaining the 5 highest loading items for each of the four main characteristics. By doing so, the items with factor loadings lower than good (λ < .55) would be dropped, while the 10 highest loading items overall would be retained. This procedure keeps the possibility to retest for the main characteristics as possible underlying constructs as well as to accurately assess our measurement as a one-dimensional superordinate construct with high inter-factor correlations among items (cf., Johnson et al., 2011).

Study 3 was conducted to further explore the factor structure of this 20-item version of the CREST in a new sample. In this study, the authors agreed on 8 selected items (i.e., 2 items of each main characteristic) to be inverted in meaning by using antonyms or adding negation. For example, “teammates maintained positive communication with each other” was reworded as “teammates started to communicate negatively with each other” and “the team gained confidence” was inverted as “the team lost its confidence”. A possible method factor was added to the exploratory models. See Appendix A for full details of the items.

A part from the CREST, the CD-RISC (Connor & Davidson, 2003) was added to the questionnaire package to test the CREST’s discriminant validity. Moreover, concurrent measures that conceptually related to team resilience characteristics, were added to provide additional evidence for the CREST’s concurrent validity. This study included the persistence subscale of the CEQS, the OCESS, and the intra-team conflict subscale of the PMCYS.

Method

Participants and procedures. A convenience sample of 345 athletes participated in the study. Both an online assessment and a paper and pencil questionnaire were employed as data collection methods. An informed consent form preceded each questionnaire package. The athletes participated in a wide variety of team sports; the majority of them (i.e., 67%) played soccer, netball, hockey or rugby. Additional demographic information can be seen in Table 1.

Measures. The 20-item version of the CREST was administered with the items randomly sorted. The same stem was used, namely “In the past month, when my team was under pressure…” Participants had to indicate to what extent they agreed or disagreed on a 7-point scale.

Thereafter, concurrent and discriminant measurements were assessed. The 5-item OCESS (Fransen, Kleinert, et al., 2014) was assessed with a 7-point scale (1 = *strongly disagree* and 2 =

strongly agree). An example item was “rate your team’s confidence that your team has the ability to encourage each other during the game”. The 5-item persistence subscale of the CEQS and the 4-items intra-team conflict subscale of the PMCYS (e.g., on this team, most athletes make negative comments that put their teammates down) was assessed likewise (Ntoumanis & Vazou, 2005; Short et al., 2005). As a discriminant measure, the 10-items CD-RISC (Connor & Davidson, 2003) was assessed with a 5-point scale (1 = *not true at all* and 5 = *true nearly all the time*). An example item was “I am able to adapt when changes occur”. An ICM CFA of these additional measures converged to a good fit (χ^2 = 329.42, *df* = 224, $p < .001$, *CFI* = .95, *TLI* = .94, *RMSEA* = .037, and *SRMR* = .053).

Results and Discussion

Similar to Study 2, different structural models were compared; fit indices and ABIC criteria were added to Table 2. With this new sample, the ICM CFA solution (CFI = .79, TLI = .75) also lacked acceptable fit indices. ESEM, both with (CFI = .99, TLI = .97) and without (CFI = .98, TLI = .97) a method factor accounting for the negative items, provided a good fit to the data. A closer examination of the standardized loadings in the ESEM including a positive and negative method factor revealed that the four main characteristics could not be retained as structural dimensions of the CREST. For the four main characteristics, factor loadings ranged from λ = -,46 (*p* = .06) to λ = .60 ($p < .001$), mean $|\lambda|$ = .18 (*SD* = .17), and cross-loadings ranged from λ = -,251 ($p < .001$) to λ = .27 ($p < .001$), mean $|\lambda|$ = .08 (*SD* = .06). Thus, similar to Study 2, it could be concluded that the four main characteristics of team resilience emerged as a one-dimensional construct. Nonetheless, the positive and the negative factor were clearly defined by their respective items (for targeted loadings, mean $|\lambda|$ = .59, *SD* = .10; and for cross-loadings, mean $|\lambda|$ = .21, *SD* = .10).

This structural difference between positively and negatively worded items was then retested within a B-ESEM to assess the extent to which the wording effect would uphold against a general team resilience factor explaining variance in all items. This B-ESEM model fitted the data well (χ^2 =

73.89, *df* = 71, *p* = .38, *CFI* = 1, *TLI* = 1, *RMSEA* = .011, and *SRMR* = .014). The resulting estimates demonstrated that no significant variance ($p < .05$) could be explained by any of the main characteristics when taking a general team resilience factor into account. In contrast, the negatively worded items still differed in structure from the positively worded items. This negative method factor (targeted loadings, mean $|\lambda| = .51$, *SD* = .07) appeared to load equally on the negatively worded items as the general resilience factor (mean $|\lambda| = .44$, *SD* = .11). All standardized factor loadings of this B-ESEM are presented in Table 4.

Subsequently, the concurrent and discriminant measures were evaluated. The ESEM-within-CFA model converged with an acceptable fit to the data ($\chi^2 = 997.75$, *df* = 727, $p < .001$, CFI = .94, *TLI* = .93, *RMSEA* = .032, and *SRMR* = .045). Significant correlations between the general resilience factor and the concurrent and discriminant measures emerged. More specifically, the G-factor, representing common variance in all CREST items, correlated significantly with the persistence subscale of the CEQS ($r = .77$, $p < .001$), the OCESS ($r = .63$, $p < .001$), and the intra-team conflicts subscale of the PMCYS ($r = -.41$, $p < .001$). The correlation of the G-factor with the CD-RISC as a discriminant measure ($r = .25$, $p < .01$) remained below the .30, i.e., the cut-off for concurrent measures.

As for the difference between the two method factors, their correlations with the persistence subscale of the CEQS and the CD-RISC were all non-significant ($p > .05$), reflecting this structural difference as a mere methodological artifact. However, the intra-team conflict subscale and the OCESS both appeared to be uncorrelated ($|r| < .15$, $p > .05$) with the unique variance among the positive items, but significantly correlated to the unique variance in the negative items ($|r| > .16$, $p < .05$). This difference in concurrence with the perceived conflicts within the team and with behavioural signs of collective efficacy could also suggest a functional difference between the positive and negative factor (Morgeson & Hofmann, 1999). Indeed, the antonyms that were preferred over negations to inverse the meaning of some items, might have partly altered their meaning. For example, “not communicating positively” might differ from “starting negative communication”. Likewise, “not gaining” confidence might not bear the same meaning as “losing” it. Such a difference between (not) demonstrating resilience and showing vulnerabilities also resonates with other sport psychology research in which positive and negative developmental trajectories have been distinguished. For example, coaches who were not supportive were not necessarily found to actively thwart their athletes’ satisfaction (e.g., Bartholomew, Ntoumanis, Ryan, & Thøgersen-Ntoumani, 2011).

To conclude, Study 3 indicated our CREST as a team resilience measure capturing both the resilient characteristics in sports teams as well as the presence of vulnerabilities under pressure. Although these positive and negative factors initially appeared as a mere methodological artefact, the factor loadings on the reverse coded items were equally high for the “vulnerabilities” factor as for the general “team resilience” factor. Moreover, a subsequent ESEM-within-CFA analysis revealed a possible functional difference between the positive and negative team characteristics as correlations with the OCESS and intra-team conflicts were only significant with the reversely coded items. In addition to Study 2, intra-team conflicts were shown to concur with (lacking) resilience. In other words, in less resilient teams, when under pressure, conflicts between players could be observed. In addition, the divergent validity of our CREST measures with the CD-RISC measure for individual resilience was supported. It appeared as our new CREST scale measured resilience at a different level focusing on the psychosocial processes in teams rather than individual athletes’ traits. However, in order to further explore the difference at individual and team level, additional data is required from whole sport teams. This additional study is also necessary to confirm the structural and functional difference between the positive and negative team resilience factor at both levels.

Study 4

This fourth and final study was conducted for confirmatory purposes. The first aim was to confirm the two-factor structure both at the individual and the team level. The second aim was to replicate evidence for the concurrent and discriminant validity of the CREST. Like in Study 3, the trait-like measure of individual resilience (CD-RISC; Connor & Davidson, 2003) was assessed as a discriminant measure to the CREST; the intra-team conflict subscales of the PMCYS (Ntoumanis & Vazou, 2005) and the OCESS (Fransen, Kleinert, et al., 2014) were assessed as concurrent measures. The third aim was to assess the CREST reliability and to quantitatively confirm the statement of Morgan et al. (2013, 2015) that team resilience is more than the sum of resilient individuals. It was hypothesized that in a multilevel mixed model the individual resilience of teammates would predict participants' perceptions of their team's resilience a part from a major (+50% of the variance) between teams effect.

Method

Participants and procedures. The board members and coaches of various sports teams were contacted to explain the purpose of this study. Upon their consent, the full team was visited before a team practice. In teams with players under the age of 16, the coaches were asked to distribute a consent form prior to our visit.

In this study, we aimed to recruit a stratified sample consisting of an equal amount of male and female teams in five different sports (i.e., basketball, handball, hockey, soccer, and volleyball) and three different age categories (i.e., U16, U19, senior) at both high and low level. We aimed at a 20:1 participant to item ratio and at least at 300 participants to reach a 70% chance of finding a complete factor structure (Stevens, 1996).

The 53 teams in our sample consisted of 8.92 players on average (*SD* = 2.19). The smallest team contained only 5 players who were willing to participate; the largest team contained 15 players. The 473 players in our sample had played 2.21 years for their current coach (*SD* = 1.73 years) with an average of 6.34 practice hours per week (*SD* = 2.36). Additional demographic information is shown in Table 1. In this sample, 216 athletes indicated their interest in participating in future studies and filled out their e-mail address. All 216 athletes were contacted via e-mail to retest the CREST online after 3 to 5 weeks. Of them, 75 (i.e., 35%) participated in the retest.

Measures. The same 20-item version of the CREST (see Appendix A) was administered through a paper and pencil questionnaire. Both players and coaches had to indicate to what extent they agreed or disagreed on a 7-point scale. Subsequently, concurrent and discriminant measurements were presented in the same way as in Study 3: the 5-items OCESS scale (Fransen, Kleinert, et al., 2014), the 4-items intra-team conflict subscale of the PMCYS (Ntoumanis & Vazou, 2005), and the 10-items CD-RISC (Connor & Davidson, 2003). Also in this study, an ICM CFA revealed a good factorial structure of these additional concepts (χ^2 = 273.92, *df* = 147, *p* < .001, *CFI* = .93, *TLI* = .92, *RMSEA* = .043, and *SRMR* = .055).

Results and Discussion

The structural invariance of the two-factor CREST

The structural difference between demonstrating resilient characteristics (positively worded items) and showing vulnerabilities under pressure (items with antonyms to the resilient characteristics) was retested with an ICM CFA approach. This confirmatory analysis of the two-factor structure appeared to fit the data well (χ^2 = 303.95, *df* = 166, *p* < .001, *CFI* = .94, *TLI* = .93, *RMSEA* = .045, and *SRMR* = .045). Furthermore, this two-factor version of the CREST showed both metric and scalar invariance across gender, level, and age categories. More specifically, no significant change in fit statistics was found when constraining factor loadings and intercepts between gender, level or age groups. Even in the rare case that the chi-square difference between nested invariance models reached a *p* < .05 level, the difference in *RMSEA* and *CFI* remained below .01. Across sports, only metric invariance could be achieved ($\Delta\chi^2$ = 72.73, Δdf = 68, *p* = .32) and (slight) scalar differences remained significant ($\Delta\chi^2$ = 111.95, Δdf = 68, *p* < .001). This metric invariance implies that the conceptual meanings of the positively and negatively worded items reflect the latent factors in the same way in each sport and language. The scalar differences imply that athletes in different sports and different countries differ in their norms to either agree or disagree with the indicators. As Lagrange multiplier tests confirmed multiple intercepts to vary over groups, no valid comparison of resilience levels can be made between different sports. In contrast, comparing team resilience levels between genders, age categories, and levels is psychometrically sound.

Invariant factor loadings could also be demonstrated between the previous sample of Study 3 in the UK and this Flemish sample. Although the chi-squares differed significantly when constraining factor loadings over samples ($\Delta\chi2 = 44.40$, $\Delta df = 18$, $p < .01$), marginal differences in RMSEA (< .001) and CFI (.005) suggested metric invariance also across different samples. Scalar differences ($\Delta\chi2 = 119.12$, $\Delta df = 18$, $p < .001$, and $\Delta CFI = .020$) between the samples were larger. It should be noted that not only the language but also the sampling method differed between samples. Participants in the UK were conveniently sampled both online and through paper and pencil questionnaires. By contrast, the Belgian participants were purposefully sampled within their teams, based on a priori stratification criteria.

The multilevel structure of the two-factor CREST

Each item was tested for how much variance could be attributed to the team level. The intra-class correlations revealed that the between team variance of items ranged from 4% to 25%, similar to the between team variance of items in concurrent scales (e.g., ICC for intra-team conflict items ranged from 9% to 19%, ICC for collective efficacy items ranged from 13% to 19%). For a multilevel CFA an ESEM-within-CFA approach was preferred. By releasing the ICM constraint that forces cross-loadings to be exactly zero, an ESEM-within-CFA approach proved to better represent real-life data (Morin et al., 2015) and to result in more reliable, less inflated, inter-factor correlations (Asparouhov & Muthén, 2009). A multilevel generalized CFA allowing for fixed between team effects revealed that in total 37% of the variance could be attributed to the team level. This 37% is a good amount of between teams variance (cf. Bliese, 2000) resulting in an acceptable reliability of group means (Spearman Brown formula with ICC = .82; Bliese, 2000; Lüdtke et al., 2009). Although this multilevel CFA demonstrated good fit to the data ($\chi^2 = 666.90$, $df = 335$, $p < .001$, $CFI = .91$, $TLI = .90$, $RMSEA = .045$, $SRMR(within) = .045$, $SRMR(between) = .13$), and data of 53 teams was obtained (Lüdtke et al., 2009), the estimated standard errors should be interpreted with caution because more parameters had to be estimated than that there were teams in our sample (Muthén & Muthén, 1998-2015).

Concurrent and discriminant validity

The correlations of the CREST scale with the concurrent measures ranged from $r = -.49$ to $r = .53$ ($p < .001$). In absolute values, the lowest correlation was observed between the positively worded factor and the intra-team conflicts ($r = -.34$, $p < .001$). In this study, the correlation between the positive en negative team resilience factors was significantly higher than their correlations with the concurrent measures ($p < .001$); notwithstanding that the concurrent measures still differently correlated with either the positive or negative CREST factor (p < .05). Individuals' resilience as a trait correlated less than .30 with the dynamic characteristics of team resilience. In absolute values, the correlations ranged from .22 ($p < .001$) to .29 ($p < .001$). Thus, similar to the previous studies, the CREST's discriminant validity was supported in comparison with the CD-RISC measure.

Reliability of measurements

For the reliability of our multilevel CFA measures, we opted for omega as a decomposed reliability coefficient (Geldhof, Preacher, & Zyphur, 2014). The CREST proved to be reliable at both the within teams level ($\omega = .90$) and the between teams level ($\omega = .99$). Furthermore, over a four-week time period, these factors appeared to be measured consistently with test-retest reliability coefficients of .69 and .70. Although a single player as informant for the team did not prove to result in a reliable rating (individual ICC = .24), the players' agreement on their team's resilience within teams ranged from 72% to 76%. The agreement between coaches and players was 75% on the positive resilience factor and 83% on the negative resilience factor.

Testing Morgan et al.'s (2013) hypothesis

Finally, two multilevel random effects models were tested to evaluate the extent to which the two team resilience factors scores reflected more than the sum of the respective teammates' individual resilience scores. Corroborating the discriminant validity of the CREST, no significant fixed effect of the individual resilience of teammates (computed as the mean of other team members reports on the CD-RISC) were revealed (for the positive factor: b = .40, p = .52; for the negative factor: b = -.48, p = .36). However, at the team level, significant random effects emerged for both the sum of the resilient individuals as well as main the team resilience level (i.e., the constant). For the positive resilience factor, the random effect of teammates resilience was .32 and the constant varied .28 points between teams. For the negative team resilience factor, the random effect of teammates resilience was 1.40 and the constant varied .38 point between teams. In these models, the total variance at team level was 53% for the resilient characteristics and 71% for the vulnerabilities under pressure.

Conclusion

To conclude, Study 4 corroborated the previous exploratory studies and provided evidence for the CREST as a measure that assesses team resilience as a two-factor construct. It could be measured as a combination of a team's ability to demonstrate resilient characteristics and the absence of vulnerabilities shown under pressure. The CREST appeared to be valid at a between-athlete and a between-team level. A multi-level ESEM-within-CFA additionally revealed that 37% of the variance could be attributed to the team level. Moreover, within teams, coaches and players seemed to agree on their assessment of their team's resilience. Finally, the CREST's concurrent and discriminant validity was confirmed; just as it was demonstrated in multilevel models that team resilience was indeed more than the sum of resilient individuals.

General Discussion

The purpose of the four consecutive studies reported in this paper was to develop and validate a measure that could assess the characteristics of resilience in sports teams. The aim was to construct a reliable referent-consensus model to quantify the current state of group-level characteristics that have been reported in qualitative research as predictors of resilient outcomes. After the clarity, relevance, and specificity of an initial pool of items were confirmed in Study 1, an optimized set of 20 items emerged from the analyses in Studies 2, 3, and 4. One specific finding was that the four main characteristics of team resilience were quantitatively represented by a single construct. This finding implies that the four main characteristics of team resilience, although separately reported in a qualitative analysis, coincide in practice. This is in line with the findings of Morgan et al. (2015) reporting underlying processes to tap into multiple characteristics and the socio-ecological perspective of Galli (2016) stressing the importance of social interactions. Indeed, it can be argued that key interactive processes like a learning approach and shared leadership concurrently affect the status of the four characteristics of team resilience. For example, when multiple leaders strategically adjust upon a teammate's injury, their team could be simultaneously provided with a plan B, re-enforced communication structures and behavioural agreements, a more prominent learning approach focusing on processing their plan B and a shared believe based on previous experiences of the particular players involved.

Yet, from the analyses of Study 3 at the individual level and Study 4 at the team level, two distinct factors emerged: (a) the team's ability to display resilient characteristics and (b) the vulnerabilities being revealed during stressors. More specifically, sport teams under pressure that were not showing vulnerabilities such as negative communication, were not necessarily learning from setbacks or demonstrating resilient characteristics such as a shared vision or exchanging social support between teammates. These two factors were evaluated as a methodological artefact, but also functional differences could be found based on the correlational structure of the data. In line with bright versus dark side psychological research in sports (e.g., Bartholomew et al., 2011), the resilient characteristics were more related to "bright" concurrent measures such as a learning approach or positive collective beliefs, while the vulnerabilities were more related to "dark" intra-team conflicts.

Although this conceptual difference was found in two different samples, it was overlooked in the interviews by Morgan et al. (2013). In the latter interviews, the resilient characteristics rather than the vulnerabilities seem to have drawn the attention of practitioners to profile their team's resilience. Integrating this knowledge in the current framework will help future studies to identify specific, and perhaps different, antecedents or consequences of each factor of team resilience.

Nonetheless, from a dynamical and process-oriented perspective, it still remains to be examined whether the positive and negative team resilience factors evolve in different stages. It could be argued that in teams developing their resilience, at first the group learns to stop displaying vulnerabilities under pressure (e.g., they stop communicating negatively). Then, in a second stage, the team positively adjusts and adapts to stressors (e.g., they learn to increase positive communication under pressure). In fact, to irrefutably link resilient characteristics to resilient outcomes, a longitudinal study is required. From the cross-sectional studies in this paper, only correlational differences suggested a functional difference between factors, but a superseding method effect could not be fully disproved.

Furthermore, all of the hypothesized concurrent and convergent measures were found to correlate correspondingly with the CREST. Higher scores on the CREST correlated positively with the persistence, effort, and collective efficacy beliefs of sport teams. Also, a peer mastery climate correlated positively with multiple team resilience characteristics. In contrast, higher values on the CREST coincided with lower values of intra-team conflicts. Moreover, the ESEM-within-CFA's demonstrated that these concurrent concepts were distinctively present in the structure of our data. In other words, these concurrent concepts, which appear to be nomologically similar, emerged as structurally different measures.

Additionally, the discriminant validity of the CREST with a measure for individual resilience was demonstrated. However, caution is warranted because the measure for individual resilience not only differs from the CREST by the referent used (individual athletes vs. the team). It also assesses resilience as a trait (vs. the state measurement of team level characteristics) and was initially developed to assess resilience in general (vs. the team sports specific items of the CREST). Nonetheless, in a multilevel regression model, it appeared that team resilience, at team level, could be computed as more than the sum of resilient individuals. Therefore, we argued that athletes' individual resilience conceptually differed from our state measure of team resilience characteristics.

From a practical point of view, when coaches mobilize strategies to enhance one of the resilient characteristics, they should also take the other characteristics into account. Indeed, in teams that succeed at developing one of the characteristics of resilience, it is likely that other characteristics may be enhanced simultaneously, or over time, in accordance with the stage of a team's development. Furthermore, relying on Morgan at al.'s (2015) finding that reflecting on the team's functioning under pressure can, by itself, enhance team resilience, we wonder whether the CREST could also be used as a self-assessment tool in sports teams. The inventory comprises of several specific behaviours and state aspects of a resilient teams. As such, they could constitute a reference point for teams to reflect on their collective functioning during stressors. More specifically, specific scenarios could be planned from individual items and peer leadership and self-regulation in these scenarios could empower the athletes to enhance their team's functioning under pressure.

Among the strengths of this multi-study paper, the combination of both qualitative and quantitative methods should firstly be highlighted. In addition to the expert panels and the cognitive interviews, multiple versions of the CREST were administered to four independent samples in two countries (Belgium and the UK). Secondly, in Study 4, whole sports teams were sampled and multi-level analyses demonstrated the usefulness of the inventory at both the between-player and the between-team level. A third strength is the possibilities the CREST has to offer for future research. With a specific measure for team level resilient characteristics, earlier conclusions of qualitative studies can now be verified. For example, it could now be examined how a team's functioning is affected by winning or losing streaks and what type of coaches' or athletes' leadership fosters the resilient characteristics of sports teams. Moreover, because the CREST has been shown to be reliable over time, it could also be used in longitudinal and experimental studies to monitor changes in team resilience.

Despite these strengths, some limitations should also be noted. First, the cross-sectional nature of the four samples in this paper prevented criterion and predictive validity of the CREST to be established. In qualitative research the characteristics of team resilience have been assumed to be predictors of resilient outcomes in sports teams. However, this assumption remains untested in this quantitative work. Future longitudinal studies are also required to investigate how specific characteristics can strengthen a sports team's resilience as a process. For instance, it still needs to be examined whether CREST measures can predict specific performance strategies employed in competition-related situations (e.g., the emotional control, focus, and tactical flexibility during different game situations).

Second, the sample in Study 2 and the retest sample in Study 4 were based on participants' self-selection. Therefore, it should be noted that the respective results might contain systematic errors due to participants' interest in group dynamics or team resilience. Nonetheless, given that the samples in Study 1 and Study 4, and the paper and pencil sample in Study 3 were targeted without self-selection, those results could be argued to be less biased by self-selection.

To conclude, we propose that the CREST inventory (as presented in Appendix A) can be used in future research to assess the characteristics of resilience in sport teams. As the two CREST factors (i.e., displaying resilient characteristics versus showing vulnerabilities) have been shown to be reliable over time, the CREST can be used to monitor both the current state and the development of resilient characteristics. Also, by conceptualising qualitative reports into a quantitative state-like measure of team level predictors of resilience, we offer scholars a tool to further investigate team resilience as a process. We also suggested that coaches could make use of the CREST as a basis for group reflections and to tailor interventions to improve their team's functioning under pressure.

References

Arnold, R., Fletcher, D., & Daniels, K. (2013). Development and validation of the Organizational Stressor Indicator for Sport Performers (OSI-SP). *Journal of Sport and Exercise Psychology, 35*(2), 180-196.

Asparouhov, T., & Muthén, B. (2009). Exploratory structural equation modeling. *Structural Equation Modeling: A Multidisciplinary Journal, 16*(3), 397-438. doi:10.1080/10705510903008204

Bartholomew, K. J., Ntoumanis, N., Ryan, R. M., & Thøgersen-Ntoumani, C. (2011). Psychological need thwarting in the sport context: Assessing the darker side of athletic experience. *Journal of Sport and Exercise Psychology, 33*(1), 75-102.

Bliese, P. D. (2000). Within-group agreement, non-independence, and reliability: Implications for data aggregation and analysis. In K. J. Klein & S. W. J. Kozlowski (Eds.), *Multilevel theory, research, and methods in organizations* (pp. 349-381). San Fransisco, CA: Jossey-Bass.

Bolter, N. D., & Weiss, M. R. (2013). Coaching behaviors and adolescent athletes' sportspersonship outcomes: Further validation of the Sportsmanship Coaching Behaviors Scale (SCBS). *Sport, Exercise, and Performance Psychology, 2*(1), 32-47. doi:10.1037/a0029802

Bonanno, G. A., Romero, S. A., & Klein, S. I. (2015). The Temporal Elements of Psychological Resilience: An Integrative Framework for the Study of Individuals, Families, and Communities. *Psychological Inquiry, 26*(2), 139-169. doi:10.1080/1047840x.2015.992677

Brown, T. A. (2015). *Confirmatory factor analysis for applied research* (2 ed.). New York: NY: Guilford Press.

Chan, D. (1998). Functional relations among constructs in the same content domain at different levels of analysis: A typology of composition models. *Journal of Applied Psychology, 83*(2), 234-246. doi:10.1037/0021-9010.83.2.234

Chan, J. C. (1991). Response-order effects in Likert-type scales. *Educational and Psychological Measurement, 51*, 531-540.

Collins, & MacNamara. (2012). The rocky road to the top: Why talent needs trauma. *Sports medicine, 42*(11), 907-914. doi:10.2165/11635140-000000000-00000

Comrey, A. L., & Lee, H. B. (1992). *A first course infactor analysis* (2 ed.). Hillsdale, NJ: Erlbaum.

Connor, K. M., & Davidson, J. R. T. (2003). Development of a new resilience scale: The Connor-Davidson Resilience Scale (CD-RISC). *Depression and Anxiety, 18*(2), 76-82. doi:10.1002/da.10113

Conrad, F. G., & Blair, J. (2004). Data quality in cognitive interviews: The case of verbal reports. In S. Presser, J. M. Rothgeb, M. P. Couper, J. T. Lessler, E. Martin, J. Martin, & E. Singer (Eds.), *Methods for testing and evaluating survey questionaires* (pp. 67-87). Hoboken, NJ: Wiley.

Cronin, L. D., & Allen, J. (2017). Development and initial validation of the Life Skills Scale for Sport. *Psychology of Sport and Exercise, 28*(1), 105-119. doi:10.1016/j.psychsport.2016.11.001

Cruickshank, A., & Collins, D. (2014). Illuminating and applying “The Dark Side”: Insights from elite team leaders. *Journal of Applied Sport Psychology, 27*(3), 249-267. doi:10.1080/10413200.2014.982771

Dietrich, H., & Ehrlenspiel, F. (2010). Cognitive interviewing: A qualitative tool for improving questionnaires in sport science. *Measurement in Physical Education and Exercise Science, 14*(1), 51-60. doi:10.1080/10913670903455025

Duda, J. L., & Hayashi, C. T. (1998). Measurement issues in cross-cultural research within sport and exercise psychology. In J. L. Duda (Ed.), *Advances in sport and exercise psychology measurement* (pp. 471-483). Morgantown, WV: FIT.

Dunn, J. G. H., Bouffard, M., & Rogers, W. T. (1999). Assessing item content-relevance in sport psychology scale-construction research: Issues and recommendations. *Measurement in Physical Education and Exercise Science, 3*(1), 15-36. doi:10.1207/s15327841mpee0301_2

Fletcher, D., Hanton, S., & Mellalieu, S. D. (2006). An organizational stress review: Conceptual and theoretical issues in competitive sport. In S. Hanton & S. D. Mellalieu (Eds.), *Literature reviews in sport psychology* (pp. 321-374). Hauppauge, NY: Nova Science.

Fletcher, D., & Sarkar, M. (2012). A grounded theory of psychological resilience in Olympic champions. *Psychology of Sport and Exercise, 13*(5), 669-678. doi:10.1016/j.psychsport.2012.04.007

Fletcher, D., & Sarkar, M. (2013). Psychological resilience: A review and critique of definitions, concepts, and theory. *European Psychologist, 18*(1), 12-23. doi:10.1027/1016-9040/A000124

Fransen, K., Kleinert, J., Dithurbide, L., Vanbeselaere, N., & Boen, F. (2014). Collective efficacy or team outcome confidence? Development and validation of the Observational Collective Efficacy Scale for Sports (OCESS). *International Journal of Sport Psychology, 45*, 121-137. doi:10.7352/IJSP 2014.45.121

Fransen, K., Vanbeselaere, N., De Cuyper, B., Vande Broek, G., & Boen, F. (2014). The myth of the team captain as principal leader: extending the athlete leadership classification within sport teams. J Sports Sci, 32(14), 1389-1397. doi:10.1080/02640414.2014.891291

Fraser-Thomas, J. L., & Cote, J. (2009). Understanding adolescents' positive and negative developmental experiences in sport. *The Sport Psychologist, 23*(1), 3-23.

Gagne, M., Ryan, R. M., & Bargmann, K. (2003). Autonomy support and need satisfaction in the motivation and well-being of gymnasts. *Journal of Applied Sport Psychology, 15*(4), 372-390. doi:10.1080/10413200390238031

Galli, N. (2016). Team Resilience. In R. J. Schinke, K. R. McGannon, & B. Smith (Eds.), *Routledge international handbook of sport psychology* (pp. 378-386). Ney York, NY: Routledge.

Galli, N., & Gonzalez, S. P. (2015). Psychological resilience in sport: A review of the literature and implications for research and practice. *International Journal of Sport and Exercise Psychology, 13*(3), 243-257. doi:10.1080/1612197x.2014.946947

Galli, N., & Vealy, R. S. (2008). "Bouncing back" from adversity: Athletes' experiences of resilience. *The Sport Psychologist, 22*, 316-335.

Garland, R. (1991). The mid-point on a rating scale: Is it desirable? *Marketing Bulletin, 2*, 66-70.

Gehlbach, H., & Brinkworth, M. E. (2011). Measure twice, cut down error: A process for enhancing the validity of survey scales. *Review of General Psychology, 15*(4), 380-387. doi:10.1037/a0025704

Geldhof, G. J., Preacher, K. J., & Zyphur, M. J. (2014). Reliability estimation in a multilevel confirmatory factor analysis framework. *Psychological Methods, 19*(1), 71-91. doi:10.1037/a0032138

Gonzalez, S. P., Moore, E. W. G., Newton, M., & Galli, N. A. (2016). Validity and reliability of the Connor-Davidson Resilience Scale (CD-RISC) in competitive sport. *Psychology of Sport and Exercise, 23,* 31-39. doi:10.1016/j.psychsport.2015.10.005

Gould, D., Dieffenbach, K., & Moffett, A. (2002). Psychological characteristics and their development in Olympic champions. *Journal of Applied Sport Psychology, 14*(3), 172-204. doi:10.1080/10413200290103482

Gucciardi, D. F., Jackson, B., Coulter, T. J., & Mallett, C. J. (2011). The Connor-Davidson Resilience Scale (CD-RISC): Dimensionality and age-related measurement invariance with Australian cricketers. *Psychology of Sport and Exercise, 12*(4), 423-433. doi:10.1016/j.psychsport.2011.02.005

Holt, N. L., & Hogg, J. M. (2002). Perceptions of stress and coping during preparations for the 1999 women's soccer World Cup finals. *The Sport Psychologist, 16*, 251-271.

Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal, 6*(1), 1-55. doi:10.1080/10705519909540118

Johnson, R. E., Rosen, C. C., & Chang, C.-H. (2011). To aggregate or not to aggregate: Steps for developing and validating higher-order multidimensional constructs. *Journal of Business and Psychology, 26*(3), 241-248. doi:10.1007/s10869-011-9238-1

Johnson, R. E., Rosen, C. C., Chang, C.-H., Djurdjevic, E., & Taing, M. U. (2012). Recommendations for improving the construct clarity of higher-order multidimensional constructs. *Human Resource Management Review, 22*(2), 62-72. doi:10.1016/j.hrmr.2011.11.006

Jowett, S., & Ntoumanis, N. (2004). The Coach–Athlete Relationship Questionnaire (CART-Q): Development and initial validation. *Scandinavian Journal of Medicine & Science in Sports, 14*, 245-257. doi:10.1046/j.1600-0838.2003.00338.x

Kampen, J. K. (2006). The impact of survey methodology and context on central tendency, nonresponse and associations of subjective indicators of government performance. *Quality & Quantity, 41*(6), 793-813. doi:10.1007/s11135-006-9026-6

Kozlowski, S. W. J. (2015). Advancing research on team process dynamics: Theoretical, methodological, and measurement considerations. *Organizational Psychology Review, 5*(4), 270-299. doi:10.1177/2041386614533586

Lee, M. J., Whitehead, J., & Ntoumanis, N. (2007). Development of the Attitudes to Moral Decision-making in Youth Sport Questionnaire (AMDYSQ). *Psychology of Sport and Exercise, 8*(3), 369-392. doi:10.1016/j.psychsport.2006.12.002

Levy, A. R., Polman, R. C. J., Nicholls, A. R., & Marchant, D. C. (2009). Sport injury rehabilitation adherence: Perspectives of recreational athletes. *International Journal of Sport and Exercise Psychology, 7*(2), 212-229. doi:10.1080/1612197x.2009.9671901

Lüdtke, O., Robitzsch, A., Trautwein, U., & Kunter, M. (2009). Assessing the impact of learning environments: How to use student ratings of classroom or school characteristics in multilevel modeling. *Contemporary Educational Psychology, 34*(2), 120-131. doi:10.1016/j.cedpsych.2008.12.001

MacKenzie, S. B., Podsakoff, P. M., & Podsakoff, N. P. (2001). Construct measurement and validation procedures in MIS and behavioral research: Integrating new and existing techniques. *MIS Quarterly, 35*(2), 293-334.

Malhotra, N., Krosnick, J. A., & Thomas, R. K. (2009). Optimal design of branching questions to measure bipolar constructs. *Public Opinion Quarterly, 73*(2), 304-324. doi:10.1093/poq/nfp023

Markland, D. (2007). The golden rule is that there are no golden rules: A commentary on Paul Barrett's recommendations for reporting model fit in structural equation modelling. *Personality and Individual Differences, 42*(5), 851-858. doi:10.1016/j.paid.2006.09.023

Marks, M. A., Mathieu, J. E., & Zaccaro, S. J. (2001). A temporally based framework and taxonomy of team processes. *Academy of Management Review, 26*(3), 356-376. doi:10.5465/amr.2001.4845785

Marsh, H. W. (2007). Application of confirmatory factor analysis and structural equation modeling in sport and exercise psychology. In G. Tenenbaum & R. C. Eklund (Eds.), *Handbook of sport psychology* (3 ed., pp. 774-798). Hoboken, NJ: Wiley.

Marsh, H. W., Morin, A. J., Parker, P. D., & Kaur, G. (2014). Exploratory structural equation modeling: an integration of the best features of exploratory and confirmatory factor analysis. *Annual Review of Clinical Psychology, 10*, 85-110. doi:10.1146/annurev-clinpsy-032813-153700

Marsh, H. W., Scalas, L. F., & Nagengast, B. (2010). Longitudinal tests of competing factor structures for the Rosenberg Self-Esteem Scale: Traits, ephemeral artifacts, and stable response styles. *Psychological Assessment, 22*, 366-381. doi:10.1037/a0019225.supp

Mellalieu, S. D., Hanton, S., & Fletcher, D. (2006). An anxiety review. In S. Hanton & S. D. Mellalieu (Eds.), *Literature reviews in sport psychology* (pp. 1-45). Hauppauge, NY: Nova Science.

Morgan, P. B. C., Fletcher, D., & Sarkar, M. (2013). Defining and characterizing team resilience in elite sport. *Psychology of Sport and Exercise, 14*(4), 549-559. doi:10.1016/j.psychsport.2013.01.004

Morgan, P. B. C., Fletcher, D., & Sarkar, M. (2015). Understanding team resilience in the world's best athletes: A case study of a rugby union World Cup winning team. *Psychology of Sport and Exercise, 16*, 91-100. doi:doi:10.1016/j.psychsport.2014.08.007

Morgeson, F. P., & Hofmann, D. A. (1999). The structure and function of collective constructs: implications for multilevel research and theory development. *Academy of Management Review, 24*(2), 249-265.

Morin, A. J. S., Arens, A. K., & Marsh, H. W. (2015). A bifactor exploratory structural equation modeling framework for the identification of distinct sources of construct-relevant psychometric multidimensionality. *Structural Equation Modeling: A Multidisciplinary Journal, 23*(1), 116-139. doi:10.1080/10705511.2014.961800

Muthén, L. K., & Muthén, B. O. (1998-2015). *Mplus user's guide*

Ntoumanis, N., & Vazou, S. (2005). Peer motivational climate in youth sport: Measurement development and validation. *Journal of Sport and Exercise Psychology, 27*, 432-455.

Pope, J. P., & Hall, C. R. (2014). Initial development of the Coach Identity Prominence Scale: A role identity model perspective. *Journal of Sport and Exercise Psychology, 36*(3), 257. doi:10.1123/jsep.2013-0039

Raykov, T. (2011). Evaluation of convergent and discriminant validity with multitrait-multimethod correlations. *British Journal of Mathematical and Statistical Psychology, 64*, 38-52. doi:10.1348/000711009X478616

Rhind, D. J. A., & Jowett, S. (2010). Relationship maintenance strategies in the coach-athlete relationship: The development of the COMPASS model. *Journal of Applied Sport Psychology*, 37-41. doi:10.1080/10413200903474472

Sarkar, M., & Fletcher, D. (2013). How should we measure psychological resilience in sport performers? *Measurement in Physical Education and Exercise Science, 17*(4), 264-280. doi:10.1080/1091367x.2013.805141

Sarkar, M., & Fletcher, D. (2014a). Ordinary magic, extraordinary performance: Psychological resilience and thriving in high achievers. *Sport, Exercise, and Performance Psychology, 3*(1), 46-60. doi:http://dx.doi.org/10.1037/spy0000003

Sarkar, M., & Fletcher, D. (2014b). Psychological resilience in sport performers: A review of stressors and protective factors. *Journal of Sports Sciences, 32*(15), 1419-1434. doi:10.1080/02640414.2014.901551

Sarkar, M., Fletcher, D., & Brown, D. J. (2015). What doesn't kill me...: Adversity-related experiences are vital in the development of superior Olympic performance. *Journal of Science and Medicine in Sport, 18*(4), 475-479. doi:10.1016/j.jsams.2014.06.010

Short, S. E., Sullivan, P., & Feltz, D. L. (2005). Development and preliminary validation of the collective efficacy questionnaire for sports. *Measurement in Physical Education and Exercise Science, 9*(3), 181-202. doi:10.1207/s15327841mpee0903_3

Sliter, K. A., & Zickar, M. J. (2014). An IRT examination of the psychometric functioning of negatively worded personality items. *Educational and Psychological Measurement, 74*(2), 214-226. doi:10.1177/0013164413504584

Stevens, J. (1996). *Applied multivariate statistics for the social sciences* (3 ed.). Mahwah, NJ: Lawrence Erlbaum Associates.

Tabei, Y., Fletcher, D., & Goodger, K. (2012). The relationship between organizational stressors and athlete burnout in soccer players. *Journal of Clinical Sport Psychology, 6*, 146-165.

Theokas, C. (2009). Youth sport participation - a view of the issues: Introduction to the special section. *Developmental Psychology, 45*(2), 303-306. doi:10.1037/a0015042

Tomas, J. M., Oliver, A., Hontangas, P. M., Sancho, P., & Galiana, L. (2015). Method effects and gender invariance of the Rosenberg Self-esteem Scale: A study on adolescents. *Acta de Investigación Psicológica, 5*(3), 2194-2203. doi:10.1016/s2007-4719(16)30009-6

Tourangeau, R., Couper, M. P., & Conrad, F. (2004). Spacing, position, and order: Interpretive heuristics for visual features of survey questions. *Public Opinion Quarterly, 68*, 368-393. doi:10.1093/poq/nfh035

Vansteenkiste, M., Mouratidis, A., & Lens, W. (2010). Detaching reasons from aims : Fair play and well-being in soccer as a function of pursuing performance-approach goals for autonomous or controlling reasons. *Journal of Sport and Exercise Psychology, 32*(2), 217-242. doi:10.1123/jsep.32.2.217

Willis, G. B. (2005). *Cognitive interviewing: A tool for improving questionnaire design*. Sage: Thousand Oaks, CA.

Yukelson, D., & Weinberg, R. (2016). Team resiliency in sport: Research to practice. In R. J. Schinke, K. R. McGannon, & B. Smith (Eds.), *Routledge international handbook of sport psychology* (pp. 547-558). New York, NY: Routledge.

Appendix A. *CREST Inventory*

The word 'team' – used in most of the questions – refers to all persons involved such as the players, coaches and team managers. When you are part of more than one team, please respond for the team that you have competed for most frequently in the past month.

The following statements describe some of the possible characteristics that sports teams may display when they experience pressure. Please keep in mind those moments that your team was under pressure in the past month and indicate the extent to which you agree or disagree with the following statements.
① represents 'strongly disagree', ⑦ represents 'strongly agree'.

In the past month, when my team was under pressure ...		strongly disagree			neutral			strongly agree
1	the team was able to focus on what was important	1	2	3	4	5	6	7
2	teammates started to communicate negatively with each other	1	2	3	4	5	6	7
3	team members fought for each other	1	2	3	4	5	6	7
4	the team lost its confidence	1	2	3	4	5	6	7
5	I felt that I could count on other members of the team	1	2	3	4	5	6	7
6	the level of collective effort in the team dropped	1	2	3	4	5	6	7
7	effective communication kept players' minds focused on the task at hand	1	2	3	4	5	6	7
8	team members started to mistrust one another	1	2	3	4	5	6	7
9	members of the team were committed to contributing to the collective belief of the team	1	2	3	4	5	6	7
10	team members fought hard to not let each other down	1	2	3	4	5	6	7
11	individuals forgot their role in the team and did not know what they had to do	1	2	3	4	5	6	7
12	the challenges we have gone through as a team helped us learn to withstand pressures	1	2	3	4	5	6	7
13	there came no support from teammates	1	2	3	4	5	6	7
14	the strong bonds between teammates helped the team during difficult times	1	2	3	4	5	6	7
15	the team could not persist through the most difficult moments	1	2	3	4	5	6	7
16	the team was able to reset their focus to alleviate pressure	1	2	3	4	5	6	7
17	the team gained belief by working together to withstand pressures	1	2	3	4	5	6	7
18	the team drew on an agreed team vision, values, and guiding behavioural principles	1	2	3	4	5	6	7
19	the team did not belief in its ability to withstand pressure	1	2	3	4	5	6	7
20	the team reflected on a shared team vision	1	2	3	4	5	6	7

Factor A (i.e., demonstrating resilient characteristics) consists of items 1, 3, 5, 7, 9, 10, 12, 14, 16, 17, 18, and 20
Factor B (i.e., vulnerabilities shown under pressure) consists of items 2, 4, 6, 8, 11, 13, 15, and 19

Table 1.

Demographic information.

	N	male	female	Age range	M age (SD)	M experience (SD)
Study 2 (in Belgium)						
players	389	189	199	16 - 52	24.0 (4.6)	14.7 (4.9)
Study 3 (in the United Kingdom)						
players	357	157	188	16 - 45	20.8 (5.17)	8.7 (7.1)
Study 4 (in Belgium)						
players	473	239	234	11 - 40	18.2 (5.0)	10.1 (5.2)
coaches	34	27	6	21 - 59	40.2 (10.2)	11.9 (9.8)

Stratification of the 53 teams in Study 4.

	High Level			Low Level			Total
	U16	U19	U21/Senior	U16	U19	U21/Senior	
Basketball	♂	♂	♂♀	♂	♂	♂♀	8
Soccer	♂♀	♂♀	♂♀	♂♀	♂♀	♂♀	12
Volleyball	♀♀	♂♀	♂♀	♂♀	♂♀	♂♀	12
Handball	♂♀	♂	♂♀	♂♀	♂♀	♀♀	11
Hockey	♂♀	♀	♂♀	♂♀	♂♀	♂	10
Total	9	7	10	9	9	9	53

Note. ♂ represents 1 male team in the sample. ♀ represents 1 female team in the sample.

Table 2.

Fit Statistics for the Structural Equation Models in Studies 2 and 3.

Model	χ^2	*df*	RMSEA [90% CI]	SRMR	CFI	TLI	ABIC
Study 2							
ICM CFA	1482.49	659	.057 [.053 - .061]	.048	.89	.88	41427
ESEM	1142.34	557	.052 [.048 - .056]	.030	.92	.90	41024
B-ESEM	899.26	524	.044 [.038 - .048]	.026	.95	.93	40938
ESEM within CFA	2264.74	1432	.039 [.036 - .042]	.037	.93	.92	-
Study 3							
ICM CFA	617.70	164	.088 [.081 - .095]	.078	.79	.75	23154
ESEM without MF	150.71	116	.029 [.013 - .041]	.025	.98	.97	22730
ESEM with MF	111.66	85	.030 [.010 - .044]	.017	.99	.97	22746
B-ESEM	73.89	71	.011 [.000 - .033]	.014	1.0	1.0	22757
ESEM within CFA	997.75	727	.032 [.027 - .037]	.045	.94	.93	-

Note. χ^2 = chi-square, *df* = degrees of freedom, RMSEA = Root Mean Square Error of Approximation, CI = Confidence Interval, SRMR = Standardized Root Mean Square Residual, CFI = Confirmatory Fit Index, TLI = Tucker-Lewis Index, ABIC = Sample-Size Adjusted Bayesian Information Criterion, ICM CFA = Independent Clusters Model in a Confirmatory Factor Analysis, ESEM = Exploratory Structural Equation Model, B-ESEM = Bifactorial Structural Equation Model, MF = method factor.

Table 3.

Standardized Factor Loadings for the Bifactorial Exploratory Structural Equation Modeling Solution of the CREST in Study 2.

Item	GS	MA	SC	CE	G-factor	Uniqueness
… effective communication kept players' minds focused	**.153**	.231**	-.006	-.029	**.664*** **	.481***
… individuals knew their role in the team and what they had to do	**.020**	.035	-.065	-.014	**.659*** **	.560***
… teammates maintained positive communication with each other	**.293**	-.076	.210	-.090	**.636*** **	.452***
… the team drew on an agreed team vision and values	**.200**	.241*	.029	-.051	**.627*** **	.505***
… the team reflected on our shared team vision	**.152**	.141	.019	.099	**.588*** **	.601***
… the team recruited and selected 'team players'	**.234**	.113	.154*	.096	**.565*** **	.581***
… communication within the team never stopped	**.263***	-.226*	-.108	.028	**.536*** **	.580***
… there was open and honest communication within the team	**.407*** **	-.103	.099	.019	**.500*** **	.563***
… the team was able to focus on what was important	.049	**-.013**	-.093	.010	**.762*** **	.408***
… the team persisted through the difficult moments	.010	**-.057**	-.126	-.279**	**.757*** **	.331***
… the challenges we have gone through as a team helped us learn	-.111	**.219***	-.145*	.116	**.733*** **	.368***
… the team was able to reset their focus to alleviate pressure	-.019	**.019**	-.202***	.047	**.730*** **	.424***
… everybody maintained a high level of collective effort	-.045	**-.036**	.247	-.264*	**.649*** **	.445***
… the team was able to move on from setbacks and not dwell on them	.027	**-.085**	-.157	-.025	**.645*** **	.551***
… the team was able to adapt to pressures due to proper preparation	.008	**.488*** **	-.121	-.115	**.632*** **	.333**
… the team was flexible in its ability to manage change	.097	**-.045**	-.170**	-.075	**.618*** **	.572***
… the team learnt positively from its experiences	.090	**.169**	-.055	.079	**.600*** **	.594***
… thorough preparation in practice helped to deal with pressures	.038	**.526*** **	-.074	-.007	**.528*** **	.438**

Table 3. *(continued)*

Item	GS	MA	SC	CE	G-factor	Uniqueness
… there was support from teammates	.045	-.075	**.150**	.208***	**.775*** **	.326***
… team members truted one another	.052	-.035	**.324** **	-.080	**.740*** **	.337***
… I felt that I could count on other members of the team	.058	-.017	**.344*** **	-.021	**.714*** **	.368***
… the strong bonds between teammates helpen during difficult times	.010	-.088	**.144**	.190*	**.708*** **	.434***
… team members fought hard to not let each other down	-.082	-.029	**.327* **	-.036	**.695*** **	.401***
… I had teammates I could rely on	.067	-.052	**.151**	.215*	**.662*** **	.485***
… I felt connected with other team members	.030	-.086	**.275**	-.002	**.642*** **	.504***
… players gave everything they could to the team	-.045	-.077	**.338**	-.191	**.641*** **	.431***
… members worked for the good of the team rather then for themselves	.053	-.122	**.065**	.214**	**.627*** **	.539***
… team members respected each other	.233	-.017	**.391* **	-.038	**.574*** **	.462***
… the team gained confidence by working together	-.053	.147*	.079	**-.147**	**.796*** **	.314***
… the team had a shared belief in its ability to withstand pressure	-.095	-.055	.009	**-.244* **	**.765*** **	.342***
… the team maintained its confidence	.015	-.028	.027	**-.282** **	**.756*** **	.347***
… members of the team were contributing to the collective belief	.057	-.125*	.115	**.196** **	**.748*** **	.370***
… going through difficult times helped the team’s collective confidence	-.234*	.118	.013	**.113**	**.731*** **	.384***
… team members fought for each other	-.260**	.191*	-.021	**.203**	**.728*** **	.324***
… we used our collective experience to increase the belief of the team	-.042	-.054	.332**	**-.103**	**.710*** **	.370***
… members of the team spread a positive belief within the team	-.050	.131	-.063	**-.011**	**.705*** **	.480***
… the team used feedback to strengthen its belief	-.006	-.155	.066	**.174**	**.689*** **	.467***

Note. GS = Group Structure. MA = Mastery Approach. SC = Social Capital. CE = Collective Efficacy.
Targeted factor loadings are in bold.
* $p < .05$. ** $p < .01$. *** $p < .001$.

Table 4.

Standardized Factor Loadings for the Bifactorial Exploratory Structural Equation Modeling Solution of the CREST in Study 3.

Item	GS	MA	SC	CE	Pos.	Neg.	G-factor	Uniqueness
Resilient characteristics								
#7	**.068**	.031	-.141	.005	**.124**	-.011	**.596*** **	.603***
#18	**.300** **	-.007	.066	-.009	**.106**	.041	**.629*** **	.497***
#20	**.510** **	-.005	.015	-.012	**.030**	.058	**.580*** **	.398**
#1	-.171**	**.164**	-.067	-.023	**.088**	-.077	**.608*** **	.555***
#12	-.002	**-.154**	.080	.017	**.040**	.093	**.579*** **	.624***
#16	.127	**.085**	-.085	.010	**.417** *	.062	**.558*** **	.481***
#5	-.101	.070	**-.078**	-.014	**.126**	-.087	**.585*** **	.613***
#10	.118	.090	**.06**	.010	**-.169**	-.097	**.688*** **	.464***
#14	-.024	-.175*	**.153**	.000	**.179**	-.066	**.723*** **	.386***
#3	-.094	.208	.096	**.043**	**.013**	.122	**.223** **	.872***
#9	.055	-.048	-.105	**-.016**	**-.368**	.027	**.655*** **	.418**
#17	.085	-.026	-.141	**-.006**	**.323**	.036	**.738*** **	.335***
Vulnerabilities under pressure								
#2	**-.105**	.085	.077	.082	-.077	**.486** **	**-.259*** **	.648***
#11	**-.021**	-.135	.134	-.079	.058	**.609*** **	**-.378*** **	.456***
#6	-.037	**-.166**	-.050	.129	.123	**.530*** **	**-.442*** **	.458***
#15	.025	**.082**	.066	.003	-.031	**.399*** **	**-.563*** **	.490***
#8	.045	.109	**-.159**	-.015	.060	**.607*** **	**-.475*** **	.248
#13	.066	.193	**.375**	-.098	.018	**.488*** **	**-.539*** **	.295**
#4	-.01	-.032	-.352	**.002**	.002	**.514*** **	**-.332*** **	-.38***
#19	-.006	-.268*	.027	**.019**	-.155	**.467*** **	**-.501*** **	.431***

Note. GS = Group Structure, MA = Mastery Approach, SC = Social Capital, CE = Collective Efficacy, Pos. = positively worded items reflecting the demonstration of resilient characteristics, Neg. = negatively worded item reflecting the vulnerabilities shown under pressure.
Targeted factor loadings are in bold.
* $p < .05$. ** $p < .01$. *** $p < .001$.