The Psychometric Properties of Two Brief Measures of Teamwork in Sport

Desmond McEwan,1,2 Eesha J. Shah,2 Kaitlin L. Crawford,2 Patricia C. Jackman,3 Matt D. Hoffmann,4 Ethan Cardinal,5 Mark W. Bruner,5 Colin D. McLaren,5 and Alex J. Benson7

1University of British Columbia, Vancouver, BC, Canada; 2University of Bath, Bath, United Kingdom; 3University of Lincoln, Lincoln, United Kingdom; 4California State University, Fullerton, CA, USA; 5Nippissing University, North Bay, ON, Canada; 6Cape Breton University, Sydney, NS, Canada; 7Western University, London, ON, Canada

In the current study, the structural and external validity of data derived from two shorter versions of the Multidimensional Assessment of Teamwork in Sport (MATS) were examined using multilevel analyses. Evidence of model–data fit was shown for both a 5-factor model comprising 19 items (with subscales assessing teamwork preparation, execution, evaluation, adjustments, and management of team maintenance) and a single-factor model comprising five items (providing a global estimate of teamwork). In general, data from both versions were positively and significantly correlated with (and distinct from) athletes’ perceptions of team cohesion, collective efficacy, performance satisfaction, enjoyment in their sport, and commitment to their team and their coaches’ transformational leadership. The measures appear well suited to detect between-teams differences, as evidenced by intra-class correlation coefficients and acceptable reliability estimates of team-level scores. In summary, the 19-item Multidimensional Assessment of Teamwork in Sport—Short and five-item Multidimensional Assessment of Teamwork in Sport—Global provide conceptually and psychometrically sound questionnaires to briefly measure teamwork in sport.

Keywords: cohesion, collective efficacy, group dynamics, leadership, performance, team effectiveness

Although the concept of teamwork has historically not received the same level of research attention in sporting contexts as it has in other team settings (e.g., health care, military, and business; Carron et al., 2012; McEwan & Beauchamp, 2014), more evidence has begun to accumulate within sport over the past decade. In their early work, McEwan and Beauchamp (2014) described teamwork as a dynamic process comprising collaborative behaviors among teammates that serve to foster the likelihood that a team will achieve its purposes. They conceptualized teamwork as a multidimensional construct composed of five higher order aspects and 14 lower order dimensions. Three dimensions comprise the “preparation” aspect—behaviors that take place prior to team tasks—including mission analysis (i.e., identifying the team’s overall purpose), goal specification (i.e., explicating performance-related goals), and planning (i.e., formulating team strategies). Three dimensions encompass teamwork “execution”—behaviors enacted during team tasks—including communication (i.e., information sharing among team members), coordination (i.e., syncing the timing and sequence of members’ interdependent actions), and cooperation (i.e., teammates working together in a unified way). Two dimensions represent the “evaluation” component—behaviors following task execution—including performance monitoring (i.e., evaluating how the team has performed) and systems monitoring (i.e., considering the various conditions that may impact the team, such as situational and environmental changes). Four dimensions comprise teamwork adjustments—behaviors enacted in response to evaluations—including problem solving (i.e., identifying potential solutions to performance-related issues), innovation (i.e., implementing novel approaches to task execution), backing up (i.e., helping teammates carry out their individual roles), and intrateam coaching (i.e., providing performance-related feedback to one another). The final two dimensions encompass the management of team maintenance (MTM)—behaviors that can appear throughout a team’s time together—including psychological support (i.e., aiding the well-being of fellow team members) and integrative conflict management (i.e., resolving conflicts within the team constructively).

Teamwork is integrated as a mediator within an input–mediator–outcome model of team effectiveness (Mathieu et al., 2008; McEwan & Beauchamp, 2014). Inputs are antecedent variables that predict team member interactions. Some of the prominent inputs that have been examined within the group dynamics in sport literature include team composition (e.g., Swaab et al., 2014), leadership (e.g., Callow et al., 2009), and organizational influences (e.g., Arnold et al., 2017). In some cases, inputs can help facilitate team functioning (e.g., effective leadership behaviors; Callow et al., 2009), whereas
in other cases, they can constrain it (e.g., organizational stress; Arnold et al., 2017). Mediators include two broad types of variables: intrateam behavioral processes (such as teamwork) as well as cognitive, affective, and motivational states that emerge as a team develops (such as team cohesion, collective efficacy, and team resilience). Those team processes and emergent states are proposed to (a) have a reciprocal relationship with one another and (b) translate inputs into outcomes, which encompass the end results that teams aim to achieve (e.g., team performance and individuals’ ongoing participation in sport). This input–mediator–outcome model can, thus, be used to understand how teamwork relates to other group and individual variables that capture team effectiveness.

Measuring Teamwork in Sport

In an effort to advance the study of teamwork, McEwan et al. (2018) developed the Multidimensional Assessment of Teamwork in Sport (MATS), a 66-item questionnaire. The MATS includes 14 sections, each of which comprises multiple items that tap into one of the 14 dimensions of teamwork. Support for some aspects of validity has been demonstrated for data derived from the MATS. For instance, to assess the structural aspect of validity—that is, the fidelity of the internal structure of data to the theoretical structure of the construct (Messick, 1995)—McEwan et al. conducted multi-level confirmatory factor analyses on data from the MATS and found that items (i.e., indicators) loaded onto the lower order dimensions specified in the conceptual model of teamwork in sport (e.g., communication and coordination); those lower order factors then loaded onto the higher order factor specified in that conceptual model (e.g., execution) at both the group (i.e., team) and individual (i.e., athlete) levels. For example, the higher order factor of “preparation” comprised the lower order factors of mission analysis (five items), goal specification (six items), and planning (six items). Evidence of internal consistency has also been demonstrated in studies using the full MATS (McEwan, 2020; McEwan et al., 2018) or certain subscales of the measure (e.g., Fransen et al., 2020; López-Gajardo, Leo, et al., 2023).

External validity assesses the degree to which data from a focal measure are related to data from other theoretically relevant constructs (Messick, 1995). This includes both convergent validity (i.e., examining whether two variables are significantly related to one another as expected) and discriminant validity (i.e., quantitative evidence supporting the distinctiveness of two variables). Since the development of the MATS, one study (to our knowledge) has examined the external validity of data derived from that 66-item measure in relation to data from other measures. Specifically, guided by meta-analytic research from other team contexts (LePine et al., 2008), McEwan (2020) found positive correlations between teamwork and associated variables that were: (a) large for athletes’ perceptions of collective efficacy, task cohesion, and satisfaction with their team’s performance; (b) moderate for social cohesion perceptions; and (c) small to medium for athletes’ enjoyment in their sport, commitment to their team, and satisfaction with their own individual performance. That study, thus, offered initial evidence of the relevance and importance of teamwork within sport.

Despite the contribution of the MATS to sport psychology research, there may be cases wherein the length of the questionnaire presents challenges to researchers and end users (e.g., applied practitioners) seeking to measure teamwork due to respondent burden (cf. Rolstad et al., 2011). For example, as noted earlier, McEwan (2020) sought to examine the links between teamwork and other variables to which research in other team settings (e.g., business and military) has demonstrated significant correlations. Due to the study’s survey containing over 100 items, the researcher split the assessment of variables into two time points in an effort to minimize participant burden—teamwork was measured at one time point, and the correlates were measured at a second point approximately 1 month later. Although teamwork was a significant predictor of various individual and group consequences (as highlighted earlier), the methodological approach limited the subsequent data analysis that could be conducted. Specifically, when conducting mediation analyses, teamwork could only be specified as the predictor variable, and as such, the study was unable to include any examination of the directionality of relationships between teamwork and emergent states (i.e., team cohesion and collective efficacy). This was acknowledged as a key limitation of the study, considering that the conceptual model of teamwork and team effectiveness explicates a reciprocal relationship between teamwork and emergent states (McEwan & Beauchamp, 2014). In other studies (e.g., Fransen et al., 2020; López-Gajardo, Leo, et al., 2023), researchers have delimitated their assessment of teamwork to a single aspect (e.g., examining teamwork execution only) to allow for teamwork to be measured alongside other variables. On the one hand, such an approach has advanced our understanding of teamwork and its relationship with other variables. For example, López-Gajardo, Leo, et al. (2023) assessed teamwork execution and team resilience at three occasions over the course of a sport season; their subsequent analysis provided evidence of a reciprocal relationship between those variables over time. On the other hand, there is an important limitation to such an approach insofar as it only allows for some components of teamwork to be tested in a study. In such cases, we are left with a partial understanding of how, and the extent to which, teamwork relates to other variables of team effectiveness.

The development of shorter versions of the MATS could help address the aforementioned limitations and would provide greater flexibility in assessments of teamwork in sport. For example, if researchers or practitioners aim to conduct a comprehensive, detailed examination of teamwork and all of its dimensions, the 66-item version could be used. Alternatively, if one only seeks to assess certain components of teamwork (e.g., the execution phase and the innovation dimension), they could include those particular subsections of the MATS in their survey pack (e.g., the 13 items that comprise the execution phase, the four items that comprise the innovation dimension). As a third option—the one that we aim to offer and evaluate in the current study—respondents could complete a shortened version of the MATS. Although this would not provide the same level of detail on each individual dimension as the “full” 66-item version of the MATS, it would, nonetheless, open up more opportunities for research in this area, such as measuring all aspects of teamwork alongside other salient variables (e.g., inputs, emergent states, and/or outcomes) on multiple occasions (e.g., in a longitudinal study). Moreover, the availability of these options could encourage more implementation research, thereby helping close the evidence-to-practice gaps that exist in many sport science disciplines, including psychology (cf. Owoseye et al., 2020).

The Current Study

There were two aims of our research. As a primary aim, we examined the structural validity of data derived from shorter versions of the MATS. Guided by the conceptual framework of teamwork by McEwan and Beauchamp (2014), as well as the validation work from the 66-item MATS (McEwan et al., 2018),
we tested the internal factor structure and reliability of data within two hypothesized factor models. The first was a five-factor model comprising 19 indicators (items) whereby all 14 dimensions and five aspects of teamwork were measured; we refer to this model and 19-item version of the MATS hereafter as the MATS-Short. The second model was a single-factor model comprising five indicators that provided content coverage of the five aspects of teamwork; as this version offers a single, global (i.e., omnibus) score of teamwork, we refer to this model and five-item version as the MATS-Global. We hypothesized that the factor structure from these two models (hypothesis 1a [H1a] and H1b, respectively) would align with the proposed conceptual structure of teamwork (see “Methods” section for details). Following best practices for designing and evaluating group-level questionnaires (Bliese et al., 2019), we also tested the extent to which the questionnaire differentiated teamwork perceptions at the between-team level and provided reliable estimates of team-level scores as well as the degree of within-team agreement.

As a secondary aim, we sought to develop the external aspect of validity. Regarding convergent validity specifically, we examined the extent to which data from both the MATS-Short and MATS-Global correlated with data from measures of three group variables—team cohesion, collective efficacy, and satisfaction with team performance—and three individual variables—enjoyment in one’s sport, commitment to one’s team, and satisfaction with one’s personal performance—that were included in previous validation work of the 66-item version of the MATS (McEwan, 2020). Comparing effect sizes of similar phenomena across studies provides insight into the robustness and magnitude of proposed relationships (beyond the statistical significance and directionality of an association). As such, we hypothesized that the correlations with teamwork would be significant, positive, and large ($r \geq 0.5$) for task cohesion (H2a), collective efficacy (H2b), and team performance satisfaction (H2c); medium ($r \geq 0.3$) for social cohesion (H2d); and small ($r \geq 0.1$) for individual performance satisfaction (H2e). As part of our assessment of external validity, we also conducted analyses of discriminant validity to ensure distinctiveness between teamwork—as measured by both the MATS-Short and MATS-Global—and each of the other group variables (task cohesion, social cohesion, collective efficacy, and team performance satisfaction).

The aforementioned correlates are all considered emergent states or outcomes of team effectiveness, which mirrors our current understanding of teamwork in sport—that is, although there is an accumulating amount of work on the consequences of teamwork, there has been relatively less research on its antecedents. Therefore, within our examination of convergent validity, we sought to build on the limited work that has studied the relationships between inputs and teamwork in sport. Specifically, we assessed the extent to which athletes’ ratings of their coaches’ transformational leadership correlated with their perceptions of teamwork. Despite the dearth of research examining teamwork in sport in relation to coach leadership, there are reasons to explore this (potential) relationship. For one, research in other team contexts has surmised that leaders play a critical role in facilitating effective teamwork (e.g., Burke et al., 2006; Mathieu et al., 2008). More specifically, transformational leadership involves behaviors that empower, inspire, and challenge followers (e.g., athletes on a team) to achieve their full potential (Bass & Riggio, 2006). A meta-analysis by Burke et al. (2006) found a medium-sized correlation between transformational leadership and team effectiveness (which included various indices, such as team productivity and perceived team performance) across a range of nonsport settings. Complementary to that work, various studies have shown that transformational leadership is, indeed, relevant and important within the context of sport (see Arthur et al., 2017 for a review). Although some initial evidence in sport points to the importance of coaches in enabling effective teamwork (e.g., Fransen et al., 2020; Sokoloff et al., 2023), more work is needed in this area to better understand how teamwork can be facilitated. Guided by the meta-analysis by Burke et al. (2006), we hypothesized that significant, medium correlations ($r \geq 0.3$) would be evident between coach transformational leadership and teamwork (H3).

**Methods**

**Participants**

Participant recruitment was guided by the estimated sample size required to conduct multilevel confirmatory factor analysis (MLCFA). MLCFA was employed in the initial MATS validation work (McEwan et al., 2018) and was deemed necessary in this study as athletes (Level 1) are nested within teams (Level 2). Although, to our knowledge, there are no definitive sample size calculations for MLCFA studies, various “rules of thumb” have been put forward, such as ensuring a minimum of 30 units at each level of analysis when individuals are nested within groups (Maas & Hox, 2005). The original sample from McEwan et al. (2018) that examined the structural validity of the 66-item MATS included 607 athletes from 48 teams. With all that in mind, we sought to obtain a heterogenous sample (in terms of age, sport, and gender) that was similar in number to that previous validation work. It should also be recognized that although McEwan (2020) provided initial evidence supporting the importance of teamwork in relation to various correlates (which were also being measured in the current study), a notable limitation of that study was that 85% of participants identified as men or boys. Therefore, as part of our purposive sampling efforts, we sought to obtain a more balanced gender ratio in this study.

The final sample comprised 674 athletes (45% women; age, mean = 21.6 ± 4.8) from 52 teams across the United Kingdom (484 athletes, 38 teams), United States of America (101 athletes, six teams), Canada (49 athletes, five teams), and Singapore (40 athletes, three teams). Sports included rugby (149 athletes, eight teams), basketball (129 athletes, 11 teams), football/soccer (74 athletes, four teams), volleyball (65 athletes, seven teams), field hockey (50 athletes, three teams), American football (41 athletes, two teams), cricket (37 athletes, three teams), lacrosse (35 athletes, two teams), netball (32 athletes, five teams), ice hockey (19 athletes, one team), water polo (14 athletes, two teams), karate (12 athletes, one team), floorball (10 athletes, one team), and curling (seven athletes, two teams). These athletes competed within a university team (508 athletes, 38 teams), at the club level (88 athletes, eight teams), as professional athletes (48 athletes, four teams), or for their country’s national team (30 athletes, two teams).

**Measures**

**Teamwork**

As discussed earlier, the 66-item MATS included 14 sections, each of which measured one of the 14 dimensions of teamwork. In each section of that version, the first item examined general perceptions of a dimension, and the remaining items tested more specific perceptions. For example, in the communication section, the first item read, “Our team communicates well with each other,”...
thereby asking participants about their overall perceptions of quality in the team’s communication; the subsequent four items asked participants to rate the amount, clarity, time efficiency, and timing of the team’s communication. For the MATS-Short, we included the first item from each of the 14 sections of the MATS (i.e., one’s general perception of a lower order teamwork dimension)—these items were pooled into five sections that represented the five higher order aspects of teamwork: preparation, execution, evaluation, adjustments, and MTM (see Table 1). For each of those five sections, we also included an item that asked participants about their overall perceptions of teamwork for that particular higher order aspect. As a result, the MATS-Short included 19 items, each of which was scored on a 7-point scale from 1 (completely disagree) to 7 (completely agree). For example, the first section included an item asking about participants’ general perceptions of their team’s preparation (“Teammates prepare for competitions together effectively”) followed by an item for each of mission analysis, goal specification, and planning (i.e., four items total).

The MATS-Global is structured in the same way as the MATS-Short but only includes the five items that tap into participants’ overall perceptions of teamwork for each of its five aspects. Both versions are available online as Supplementary Materials S1 and S2 (available online) and on the project’s Open Science Framework (OSF) page (https://osf.io/pqnsw/). By testing two-factor models (detailed later), our goal was to maximize flexibility in future teamwork studies whereby researchers could use (a) the 66-item MATS to obtain a detailed and comprehensive assessment of teamwork, (b) the 19-item MATS-Short to measure all 14 dimensions and obtain estimates of the five aspects of teamwork in a less detailed but more time-efficient manner, or (c) the five-item MATS-Global to obtain a general “snapshot” of teamwork based on participants’ overall perceptions of the five aspects of teamwork.

Correlates of Teamwork

The 18-item Group Environment Questionnaire (Carron et al., 1985) was used to assess team cohesion. Items were scored on a 9-point scale, from 1 (strongly disagree) to 9 (strongly agree) for both task (e.g., “Our team is united in trying to reach its goals for performance”) and social (e.g., “Members of our team stick together outside of practices and games”) cohesion. Support has been demonstrated for the validity and reliability of data derived from the Group Environment Questionnaire (Carron et al., 1985). Although the Group Environment Questionnaire can estimate four dimensions of cohesion—that is, (a) individuals’ attractions to the group’s task objectives and (b) social interactions, as well as their perceptions about the group’s overall unity around its (c) task objectives and (d) social connections—cohesion has often been measured as “social cohesion” and “task cohesion” such as in the meta-analysis of cohesion and performance by Carron et al. (2002). The two-factor approach to studying cohesion has also been used in several primary studies, including in the initial work examining its relationships with teamwork in sport (McEwan, 2020). We, therefore, examined cohesion as task cohesion and social cohesion. In the current study, McDonald’s omega was 0.88 for task cohesion and 0.90 for social cohesion.

Collective efficacy was measured with the four-item “Ability” subscale from the Collective Efficacy Questionnaire for Sports (Short et al., 2005). For this subscale, participants were asked to rate their confidence in their team’s ability to outperform opposing teams (e.g., “to play more skillfully than the opponent”) in an upcoming competition from 1 (not at all confident) to 10 (extremely confident). Short et al. (2005) provided evidence of validity and reliability for data derived from this questionnaire. In the current study, McDonald’s omega was 0.96.

The Sport Commitment Questionnaire (Scanlan et al., 1993) was used to measure athletes’ enjoyment in their sport and commitment to their team. The four-item “Sport Enjoyment” subscale included a 5-point scale from 1 (not at all) to 5 (very much) to assess enjoyment (e.g., “Do you enjoy playing your sport this season?”). The four-item “Sport Commitment” subscale used a 5-point scale, with three items (e.g., “How dedicated are you to this team?”) measured from 1 (not at all) to 5 (very much) and one item (“What would you be willing to do to keep playing with this team?”) measured from 1 (nothing at all) to 5 (a lot of things). Support for the reliability and validity of data derived from both subscales has been shown in previous research (Scanlan et al., 1993). In the current study, McDonald’s omega was 0.96 for enjoyment and 0.89 for commitment.

The Athlete Satisfaction Questionnaire (Reimer & Chelladurai, 1998) was used to measure satisfaction with performance. Specifically, participants were asked to indicate their satisfaction with their team’s performance and with their individual performance in the current competitive season using the three-item “Team” subscale (e.g., “The team’s overall performance this season”) and the three-item “Individual” subscale (e.g., “The degree to which I have achieved my performance goals during the season”), respectively. Items within both subscales were assessed on a 7-point scale from 1 (not at all satisfied) to 7 (extremely satisfied). Evidence supporting data validity and reliability has been found in previous research employing this questionnaire (Reimer & Chelladurai, 1998). In the current study, McDonald’s omega was 0.91 for team performance satisfaction and 0.87 for individual performance satisfaction.

To measure athletes’ ratings of their coaches’ transformational leadership, we adapted the Transformational Teaching Questionnaire (Beauchamp et al., 2010). As the questionnaire was originally developed to assess transformational leadership in teaching settings (e.g., physical education), we changed the stem on the questionnaire from “my teacher” to “my coach” as well as any references to “students” and “class” within individual items to “team members” and “team” (respectively). In this measure, respondents were asked to rate the extent to which their leader demonstrated transformational leadership across 16 items (e.g., “My coach behaves as someone I can trust”) on a 5-point scale from 0 (not at all) to 5 (frequently). Support for the reliability and validity of data has been evidenced in previous studies using this measure (Beauchamp et al., 2010). In the current study, McDonald’s omega was 0.96.

Procedure

Following institutional ethics board approval, competitive, adult-aged sport teams were recruited to participate in the study by contacting team coaches. For feasibility reasons, teams were able to participate at any point in the season so long as it occurred at least 1 month after the season’s commencement and the team had competed in at least two competitions/games. When coaches expressed interest in having their team participate, researchers met with the team’s players on a single occasion before or after a team training/practice. At the data collection session, those who provided written consent to participate in the study were asked to anonymously complete the survey package, which comprised the MATS-Short and all external variables. The data collection session typically lasted 10–15 min. The full survey and the deidentified data set are provided on the project’s OSF page (https://osf.io/pqnsw/).
Table 1  Items, Descriptive Statistics, $r_{WG}$ Values, ICCs, and Factor Loadings for the 19-Item MATS-Short and 5-Item MATS-Global

<table>
<thead>
<tr>
<th>Subscale</th>
<th>Mean (SD)</th>
<th>$r_{WG}$</th>
<th>ICC(1)</th>
<th>ICC(2)</th>
<th>Within $\beta$ (SE)</th>
<th>Between $\beta$ (SE)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Preparation subscale</strong> ($M = 5.58, SD = 0.99$)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Teammates prepare for competitions together effectively</td>
<td>5.55 (1.09)</td>
<td>.80</td>
<td>.27</td>
<td>.83</td>
<td>0.80 (0.02)</td>
<td>1.00 (0.00)</td>
</tr>
<tr>
<td>Our team has identified an overall purpose for being together</td>
<td>5.85 (1.10)</td>
<td>.79</td>
<td>.17</td>
<td>.73</td>
<td>0.76 (0.02)</td>
<td>1.00 (0.00)</td>
</tr>
<tr>
<td>Our team identifies specific team goals in order to achieve our overall purpose(s)</td>
<td>5.65 (1.21)</td>
<td>.74</td>
<td>.22</td>
<td>.79</td>
<td>0.82 (0.02)</td>
<td>0.72 (0.08)</td>
</tr>
<tr>
<td>We make action plans for how we will achieve our team goals</td>
<td>5.30 (1.39)</td>
<td>.67</td>
<td>.26</td>
<td>.82</td>
<td>0.76 (0.02)</td>
<td>0.70 (0.10)</td>
</tr>
<tr>
<td><strong>Execution subscale</strong> ($M = 5.52, SD = 1.05$)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Our team demonstrates a high level of teamwork while competing against other teams</td>
<td>5.67 (1.14)</td>
<td>.76</td>
<td>.22</td>
<td>.79</td>
<td>0.80 (0.02)</td>
<td>0.93 (0.05)</td>
</tr>
<tr>
<td>Overall, team members coordinate actions well with each other</td>
<td>5.43 (1.16)</td>
<td>.78</td>
<td>.21</td>
<td>.78</td>
<td>0.87 (0.02)</td>
<td>0.99 (0.05)</td>
</tr>
<tr>
<td>In general, teammates work together effectively</td>
<td>5.64 (1.13)</td>
<td>.79</td>
<td>.23</td>
<td>.79</td>
<td>0.90 (0.02)</td>
<td>1.00 (0.00)</td>
</tr>
<tr>
<td>Our team communicates well with each other</td>
<td>5.32 (1.28)</td>
<td>.73</td>
<td>.23</td>
<td>.79</td>
<td>0.81 (0.02)</td>
<td>0.98 (0.05)</td>
</tr>
<tr>
<td><strong>Evaluation subscale</strong> ($M = 5.39, SD = 1.25$)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>After competitions, we reflect on our team’s execution</td>
<td>5.50 (1.45)</td>
<td>.68</td>
<td>.29</td>
<td>.84</td>
<td>0.79 (0.02)</td>
<td>0.88 (0.06)</td>
</tr>
<tr>
<td>Our team monitors its performance</td>
<td>5.47 (1.35)</td>
<td>.69</td>
<td>.33</td>
<td>.87</td>
<td>0.85 (0.02)</td>
<td>0.92 (0.05)</td>
</tr>
<tr>
<td>Our team monitors any factors that may affect us</td>
<td>5.20 (1.43)</td>
<td>.60</td>
<td>.26</td>
<td>.82</td>
<td>0.83 (0.02)</td>
<td>1.00 (0.00)</td>
</tr>
<tr>
<td><strong>Adjustments subscale</strong> ($M = 5.38, SD = 1.06$)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>On this team, we work together to improve between competitions</td>
<td>5.76 (1.18)</td>
<td>.78</td>
<td>.26</td>
<td>.82</td>
<td>0.81 (0.02)</td>
<td>0.97 (0.05)</td>
</tr>
<tr>
<td>Our team problem solves when we have not performed well</td>
<td>5.23 (1.33)</td>
<td>.69</td>
<td>.21</td>
<td>.77</td>
<td>0.80 (0.02)</td>
<td>1.00 (0.00)</td>
</tr>
<tr>
<td>Our team modifies our approaches when necessary</td>
<td>5.41 (1.25)</td>
<td>.71</td>
<td>.15</td>
<td>.68</td>
<td>0.78 (0.02)</td>
<td>0.86 (0.07)</td>
</tr>
<tr>
<td>Team members provide verbal feedback to each other about how to improve their individual performance</td>
<td>5.31 (1.41)</td>
<td>.63</td>
<td>.14</td>
<td>.68</td>
<td>0.80 (0.02)</td>
<td>0.92 (0.04)</td>
</tr>
<tr>
<td>Teammates demonstrate to each other how they can improve their personal performance</td>
<td>5.17 (1.55)</td>
<td>.65</td>
<td>.16</td>
<td>.72</td>
<td>0.80 (0.02)</td>
<td>0.86 (0.05)</td>
</tr>
<tr>
<td><strong>MTM subscale</strong> ($M = 5.57, SD = 1.21$)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Our team works together to ensure that members’ issues do not prevent us from being successful</td>
<td>5.29 (1.36)</td>
<td>.71</td>
<td>.21</td>
<td>.77</td>
<td>0.88 (0.02)</td>
<td>1.00 (0.00)</td>
</tr>
<tr>
<td>Teammates resolve conflicts with each other effectively if they arise</td>
<td>5.21 (1.50)</td>
<td>.71</td>
<td>.17</td>
<td>.73</td>
<td>0.78 (0.02)</td>
<td>0.80 (0.09)</td>
</tr>
<tr>
<td>Members provide support to teammates who are experiencing personal struggles</td>
<td>5.61 (1.39)</td>
<td>.60</td>
<td>.26</td>
<td>.82</td>
<td>0.70 (0.03)</td>
<td>0.88 (0.06)</td>
</tr>
<tr>
<td><strong>Five-item MATS-Global</strong> ($M = 5.56, SD = 0.94$)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Teammates prepare for competitions together effectively</td>
<td>5.55 (1.09)</td>
<td>.80</td>
<td>.27</td>
<td>.83</td>
<td>0.68 (0.03)</td>
<td>0.97 (0.05)</td>
</tr>
<tr>
<td>Our team demonstrates a high level of teamwork while competing against other teams</td>
<td>5.67 (1.14)</td>
<td>.76</td>
<td>.22</td>
<td>.79</td>
<td>0.68 (0.03)</td>
<td>0.76 (0.09)</td>
</tr>
<tr>
<td>After competitions, we reflect on our team’s execution</td>
<td>5.50 (1.45)</td>
<td>.68</td>
<td>.29</td>
<td>.84</td>
<td>0.59 (0.03)</td>
<td>0.66 (0.08)</td>
</tr>
<tr>
<td>On this team, we work together to improve between competitions</td>
<td>5.76 (1.18)</td>
<td>.78</td>
<td>.26</td>
<td>.82</td>
<td>0.80 (0.03)</td>
<td>0.91 (0.06)</td>
</tr>
<tr>
<td>Our team works together to ensure that members’ issues do not prevent us from being successful</td>
<td>5.29 (1.36)</td>
<td>.71</td>
<td>.21</td>
<td>.77</td>
<td>0.63 (0.03)</td>
<td>0.85 (0.07)</td>
</tr>
</tbody>
</table>

Note. $r_{WG} =$ median score of within-group agreement; ICC(1) = intraclass correlations denoting the amount of variance in each item that was observed at the group level; ICC(2) = intraclass correlations estimating the reliability of the group-level means for each item; $\beta =$ standardized parameter estimate of item onto factor; SE = standard error; MTM = management of team maintenance; MATS = Multidimensional Assessment of Teamwork in Sport. All beta values are significant at $p < .001$. The original MATS-Short model produced negative residual variances that led to standardized factor loadings $>$1. As such, we constrained these residual variances to 0, which did not substantially alter model fit. Therefore, the values from this model are provided in this table (please see the Supplementary Material S3 [available online] for the original model).
Data Analyses

Structural Validity

Similar to the validation work for the original MATS (McEwan et al., 2018), we conducted ML-CFA to examine the factor structure of data. In our hypothesized model for the 19-item MATS-Short (testing H1a), individual items were specified (at both the athlete and team levels) to load onto a latent factor that aligned with the conceptual model by McEwan and Beauchamp (2014). Specifically, Items 1–4 loaded onto a factor that we labeled “preparation” (Factor 1), Items 5–8 loaded onto a factor labeled “execution” (Factor 2), Items 9–11 loaded onto a factor labeled “evaluation” (Factor 3), Items 12–16 loaded onto a factor labeled “adjustments” (Factor 4), and Items 17–19 loaded onto a factor labeled “MTM” (factor 5). In our hypothesized model for the MATS-Global (H1b), we delimited the model to only include the five items that tapped into participants’ overall perceptions of the five aspects of teamwork (Items 1, 5, 9, 12, and 17). Those indicators were all specified to load onto a single factor (at both the athlete and team levels) labeled “teamwork.” A summary of both models is provided in Table 1.

All of the aforementioned factor analyses were conducted in Mplus (version 8.9; Muthén & Muthén, 2021). The codebook for these analyses is available on the project’s OSF page (https://osf.io/pqnsw/). The data were specified as ordinal, and the weighted least squares means and variance-adjusted estimation method was employed for all models as this method is robust when conducting multilevel factor analyses with a large (e.g., over 200 participants) sample (Brown, 2006; Muthén et al., 2015). Absolute data—model fit was evaluated using standardized root square mean residual (SRMR) at both the within and between levels as well as the root mean square error of approximation (RMSEA). The chi-square test of model fit was also considered and is reported for descriptive purposes; however, it has been suggested that alternative indices (such as SRMR and RMSEA) should be relied upon over the chi-square test to assess global fit as obtaining a nonsignificant chi-square value (which would provide evidence of global fit) is highly unlikely (Brown, 2006). Comparative fit was assessed using the comparative fit index (CFI) and Tucker–Lewis index (TLI). Evidence of global fit of the hypothesized model and resulting data is indicated when SRMR and RMSEA values are below .08, whereas evidence of comparative fit is suggested when CFI and TLI values exceed .90 (Brown, 2006; Hu & Bentler, 1995; MacCallum et al., 1996). Localized (mis)fit was evaluated by scrutinizing items whose residual variances were above 2.5 or factor loadings were below 0.60 while also weighing the potential consequences of removing an item on content validity (Brown, 2006; Flora & Flake, 2017).

Reliability

The internal consistency of data was assessed with two within-level criteria and two between-level criteria: athlete-level Cronbach’s alpha (α), athlete-level McDonald’s omega (Ω), team-level ordinal composite reliability (CR), and team-level average variance extracted (AVE). Support for data reliability is evident if α and Ω exceed .70 (Nunnally & Bernstein, 1994), CR exceeds .70 (Zumbo et al., 2007), and AVE exceeds .50 (Fornell & Larcker, 1981).

Group-Level Properties of Measure

We computed three indices to gain insight into the extent to which items were optimized to detect between-team differences, using both Mplus (Muthén & Muthén, 2021) and the “multilevel” package in R (Bliese, 2022). Item-level intraclass correlations were computed to estimate the amount of variance that was observed at the group level, referred to as ICC(1), as well as the reliability of the group-level means, referred to as ICC(2). As Bliese et al. (2019) noted in relation to optimizing the measurement of group constructs, higher ICC(1) values evidence a scale’s ability to differentiate between teams and, thus, augment predictive validity. ICC(2) values should ideally exceed .70 (Jebb et al., 2019). We also computed an index of within-group agreement ($r_{wg}$). LeBreton and Senter (2008) suggested that $r_{wg}$ values can be interpreted as indicating a lack of agreement (.00–.30), weak agreement (.31–.50), moderate agreement (.51–.70), or strong agreement (.71–.90).

External Validity

Multilevel correlations were carried out to assess convergent validity—that is, the relationships between teamwork and correlates at both the athlete (i.e., within-group) and team (i.e., between-group) level (H2 and H3). For the 19-item MATS-Short, teamwork was represented by the five factors/aspects of teamwork (preparation, execution, evaluation, adjustments, and MTM). For the five-item MATS-Global, teamwork was represented as a general/omnibus score. Small-, medium-, and large-sized correlations are evident when standardized beta values exceed .1, .3, and .5 (respectively; Cohen, 1992). The results of these analyses are provided in Table 2. To assess discriminant validity, we calculated the heterotrait–monotrait ratio of correlations (HTMT) values between teamwork and the group constructs (i.e., task cohesion, social cohesion, collective efficacy, and team performance satisfaction) that were measured. HTMT examines the similarity between latent variables (Henseler et al., 2015)—this assessment is important to ensure that different measures are, indeed, tapping into distinct constructs. Evidence of discriminant validity is offered when HTMT is smaller than .85.

Results

Structural Validity

MATS-Short

In our first hypothesized model, each of the 19 items was specified to load onto one of five factors, which represented an aspect of teamwork. The chi-square test of model fit (949.14) was significant ($p < .001$). Global fit was evident based on RMSEA (.06), SRMR-within (.05), and SRMR-between (.09) values. Comparative fit was also acceptable based on CFI (.92) and TLI (.91) values. All factor loadings exceeded 0.60, and all residual variances were below 2.5, thereby suggesting no concerning areas of localized misfit.

MATS-Global

In our second hypothesized model, the five indicators assessing general perceptions of the five aspects of teamwork were specified to load onto a single factor. The chi-square test of model fit (61.21) was significant ($p < .001$). Global fit was evident based on SRMR-within (.05) and SRMR-between (.05); RMSEA (.087) fell just short of the recommended .08 cutoff for acceptable fit and was, therefore, considered “borderline” (MacCallum et al., 1996). Comparative fit was acceptable as CFI = .96 and TLI = .91. There did not appear to be any substantive areas of localized misfit; all standardized factor loadings were above 0.60 except for one that fell just below that recommended cutoff (within-level evaluation = 0.59). Moreover, all residual variances were below 2.5.
Table 2  Athlete- and Team-Level Relationships Between Teamwork and External Variables

<table>
<thead>
<tr>
<th>Variable</th>
<th>Task cohesion</th>
<th>Social cohesion</th>
<th>Collective efficacy</th>
<th>Team performance satisfaction</th>
<th>Individual performance satisfaction</th>
<th>Enjoyment</th>
<th>Commitment</th>
<th>Coach transformational leadership</th>
</tr>
</thead>
<tbody>
<tr>
<td>Athlete level</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Teamwork&lt;sup&gt;a&lt;/sup&gt;</td>
<td>0.65 (0.02)&lt;sup&gt;***&lt;/sup&gt;</td>
<td>0.41 (0.02)&lt;sup&gt;***&lt;/sup&gt;</td>
<td>0.43 (0.02)&lt;sup&gt;***&lt;/sup&gt;</td>
<td>0.33 (0.03)&lt;sup&gt;***&lt;/sup&gt;</td>
<td>0.29 (0.02)&lt;sup&gt;***&lt;/sup&gt;</td>
<td>0.37 (0.02)&lt;sup&gt;***&lt;/sup&gt;</td>
<td>0.33 (0.02)&lt;sup&gt;***&lt;/sup&gt;</td>
<td>0.34 (0.02)&lt;sup&gt;***&lt;/sup&gt;</td>
</tr>
<tr>
<td>Preparation&lt;sup&gt;b&lt;/sup&gt;</td>
<td>0.53 (0.02)&lt;sup&gt;***&lt;/sup&gt;</td>
<td>0.34 (0.02)&lt;sup&gt;***&lt;/sup&gt;</td>
<td>0.35 (0.03)&lt;sup&gt;***&lt;/sup&gt;</td>
<td>0.30 (0.04)&lt;sup&gt;***&lt;/sup&gt;</td>
<td>0.28 (0.03)&lt;sup&gt;***&lt;/sup&gt;</td>
<td>0.37 (0.02)&lt;sup&gt;***&lt;/sup&gt;</td>
<td>0.34 (0.02)&lt;sup&gt;***&lt;/sup&gt;</td>
<td>0.33 (0.02)&lt;sup&gt;***&lt;/sup&gt;</td>
</tr>
<tr>
<td>Execution&lt;sup&gt;b&lt;/sup&gt;</td>
<td>0.58 (0.02)&lt;sup&gt;***&lt;/sup&gt;</td>
<td>0.38 (0.02)&lt;sup&gt;***&lt;/sup&gt;</td>
<td>0.40 (0.03)&lt;sup&gt;***&lt;/sup&gt;</td>
<td>0.31 (0.04)&lt;sup&gt;***&lt;/sup&gt;</td>
<td>0.29 (0.02)&lt;sup&gt;***&lt;/sup&gt;</td>
<td>0.35 (0.02)&lt;sup&gt;***&lt;/sup&gt;</td>
<td>0.31 (0.03)&lt;sup&gt;***&lt;/sup&gt;</td>
<td>0.31 (0.02)&lt;sup&gt;***&lt;/sup&gt;</td>
</tr>
<tr>
<td>Evaluation&lt;sup&gt;b&lt;/sup&gt;</td>
<td>0.43 (0.03)&lt;sup&gt;***&lt;/sup&gt;</td>
<td>0.28 (0.03)&lt;sup&gt;***&lt;/sup&gt;</td>
<td>0.30 (0.02)&lt;sup&gt;***&lt;/sup&gt;</td>
<td>0.18 (0.04)&lt;sup&gt;***&lt;/sup&gt;</td>
<td>0.20 (0.03)&lt;sup&gt;***&lt;/sup&gt;</td>
<td>0.22 (0.04)&lt;sup&gt;***&lt;/sup&gt;</td>
<td>0.17 (0.04)&lt;sup&gt;***&lt;/sup&gt;</td>
<td>0.24 (0.04)&lt;sup&gt;***&lt;/sup&gt;</td>
</tr>
<tr>
<td>Adjustments&lt;sup&gt;b&lt;/sup&gt;</td>
<td>0.63 (0.02)&lt;sup&gt;***&lt;/sup&gt;</td>
<td>0.39 (0.02)&lt;sup&gt;***&lt;/sup&gt;</td>
<td>0.42 (0.02)&lt;sup&gt;***&lt;/sup&gt;</td>
<td>0.29 (0.04)&lt;sup&gt;***&lt;/sup&gt;</td>
<td>0.27 (0.02)&lt;sup&gt;***&lt;/sup&gt;</td>
<td>0.34 (0.03)&lt;sup&gt;***&lt;/sup&gt;</td>
<td>0.29 (0.03)&lt;sup&gt;***&lt;/sup&gt;</td>
<td>0.36 (0.03)&lt;sup&gt;***&lt;/sup&gt;</td>
</tr>
<tr>
<td>MTM&lt;sup&gt;b&lt;/sup&gt;</td>
<td>0.60 (0.03)&lt;sup&gt;***&lt;/sup&gt;</td>
<td>0.43 (0.02)&lt;sup&gt;***&lt;/sup&gt;</td>
<td>0.38 (0.03)&lt;sup&gt;***&lt;/sup&gt;</td>
<td>0.27 (0.03)&lt;sup&gt;***&lt;/sup&gt;</td>
<td>0.24 (0.03)&lt;sup&gt;***&lt;/sup&gt;</td>
<td>0.32 (0.03)&lt;sup&gt;***&lt;/sup&gt;</td>
<td>0.27 (0.02)&lt;sup&gt;***&lt;/sup&gt;</td>
<td>0.29 (0.04)&lt;sup&gt;***&lt;/sup&gt;</td>
</tr>
<tr>
<td>Team level</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Teamwork&lt;sup&gt;a&lt;/sup&gt;</td>
<td>0.88 (0.04)&lt;sup&gt;***&lt;/sup&gt;</td>
<td>0.48 (0.10)&lt;sup&gt;***&lt;/sup&gt;</td>
<td>0.63 (0.10)&lt;sup&gt;***&lt;/sup&gt;</td>
<td>0.31 (0.18)</td>
<td>0.71 (0.14)&lt;sup&gt;***&lt;/sup&gt;</td>
<td>0.52 (0.12)&lt;sup&gt;***&lt;/sup&gt;</td>
<td>0.75 (0.04)&lt;sup&gt;***&lt;/sup&gt;</td>
<td>0.59 (0.08)&lt;sup&gt;***&lt;/sup&gt;</td>
</tr>
<tr>
<td>Preparation&lt;sup&gt;b&lt;/sup&gt;</td>
<td>0.69 (0.08)&lt;sup&gt;***&lt;/sup&gt;</td>
<td>0.40 (0.10)&lt;sup&gt;***&lt;/sup&gt;</td>
<td>0.55 (0.12)&lt;sup&gt;***&lt;/sup&gt;</td>
<td>0.25 (0.19)</td>
<td>0.67 (0.18)&lt;sup&gt;***&lt;/sup&gt;</td>
<td>0.19 (0.14)</td>
<td>0.69 (0.11)&lt;sup&gt;***&lt;/sup&gt;</td>
<td>0.54 (0.08)&lt;sup&gt;***&lt;/sup&gt;</td>
</tr>
<tr>
<td>Execution&lt;sup&gt;b&lt;/sup&gt;</td>
<td>0.95 (0.03)&lt;sup&gt;***&lt;/sup&gt;</td>
<td>0.49 (0.09)&lt;sup&gt;***&lt;/sup&gt;</td>
<td>0.59 (0.10)&lt;sup&gt;***&lt;/sup&gt;</td>
<td>0.42 (0.16)&lt;sup&gt;**&lt;/sup&gt;</td>
<td>0.72 (0.14)&lt;sup&gt;***&lt;/sup&gt;</td>
<td>0.72 (0.09)&lt;sup&gt;***&lt;/sup&gt;</td>
<td>0.62 (0.11)&lt;sup&gt;***&lt;/sup&gt;</td>
<td>0.58 (0.09)&lt;sup&gt;***&lt;/sup&gt;</td>
</tr>
<tr>
<td>Evaluation&lt;sup&gt;b&lt;/sup&gt;</td>
<td>0.62 (0.10)&lt;sup&gt;***&lt;/sup&gt;</td>
<td>0.36 (0.09)&lt;sup&gt;***&lt;/sup&gt;</td>
<td>0.58 (0.11)&lt;sup&gt;***&lt;/sup&gt;</td>
<td>0.22 (0.17)</td>
<td>0.53 (0.15)&lt;sup&gt;**&lt;/sup&gt;</td>
<td>0.17 (0.12)</td>
<td>0.66 (0.11)&lt;sup&gt;***&lt;/sup&gt;</td>
<td>0.47 (0.10)&lt;sup&gt;***&lt;/sup&gt;</td>
</tr>
<tr>
<td>Adjustments&lt;sup&gt;b&lt;/sup&gt;</td>
<td>0.80 (0.07)&lt;sup&gt;***&lt;/sup&gt;</td>
<td>0.35 (0.07)&lt;sup&gt;***&lt;/sup&gt;</td>
<td>0.63 (0.13)&lt;sup&gt;***&lt;/sup&gt;</td>
<td>0.30 (0.19)</td>
<td>0.69 (0.14)&lt;sup&gt;***&lt;/sup&gt;</td>
<td>0.60 (0.14)&lt;sup&gt;***&lt;/sup&gt;</td>
<td>0.68 (0.13)&lt;sup&gt;***&lt;/sup&gt;</td>
<td>0.63 (0.10)&lt;sup&gt;***&lt;/sup&gt;</td>
</tr>
<tr>
<td>MTM&lt;sup&gt;b&lt;/sup&gt;</td>
<td>0.88 (0.05)&lt;sup&gt;***&lt;/sup&gt;</td>
<td>0.53 (0.10)&lt;sup&gt;***&lt;/sup&gt;</td>
<td>0.51 (0.12)&lt;sup&gt;***&lt;/sup&gt;</td>
<td>0.20 (0.19)</td>
<td>0.52 (0.16)&lt;sup&gt;***&lt;/sup&gt;</td>
<td>0.66 (0.11)&lt;sup&gt;***&lt;/sup&gt;</td>
<td>0.67 (0.10)&lt;sup&gt;***&lt;/sup&gt;</td>
<td>0.47 (0.10)&lt;sup&gt;***&lt;/sup&gt;</td>
</tr>
</tbody>
</table>

Note. MTM = management of team maintenance. These values represent standardized regression coefficients (with standard errors in brackets). To ease estimation, latent decomposition in Mplus was used to decompose variance within teams (i.e., athlete level) from between teams (i.e., team level) using the observed subscale scores.

<sup>a</sup>Teamwork” is the mean score derived from the 5-item MATS-Global. <sup>b</sup>Preparation,” “Execution,” “Evaluation,” “Adjustments,” and “MTM” are subscale scores from the 19-item MATS-Short. 

**p < .01. ***p < .001.
Overall, these results can be interpreted as evidence supporting the structural validity of data within both the 19-item, five-factor model (H1a) and the five-item, single-factor model (H1b; see Table 1). It should be noted that various alternative models were also tested, as per guidance from Flora and Flake (2017). These alternative models included single-level CFAs for both versions of the measure (wherein the nesting of the data was not accounted for) as well as single-factor models (at the athlete level, team level, or both levels) for the 19-item version. In all cases, data—model fit was poorer compared with the two hypothesized models. These results provided further support for conducting multilevel CFAs (particularly when also considering the item-level ICCs) and for specifying multiple factors in the MATS-Short model.

**Reliability and Group-Level Properties**

With regard to the 19-item MATS-Short, $\alpha = .85$, $\Omega = .85$, CR = .95, and AVE = .84 for the preparation subscale; for execution, $\alpha = .91$, $\Omega = .91$, CR = .99, and AVE = .95; for evaluation, $\alpha = .87$, $\Omega = .87$, CR = .99, and AVE = .96; for adjustments, $\alpha = .87$, $\Omega = .86$, CR = .98, and AVE = .88; and for MTM, $\alpha = .82$, $\Omega = .83$, CR = .96, and AVE = .90. For the five-item MATS-Global, $\alpha = .81$, $\Omega = .81$, CR = .92, and AVE = .72. These results suggested evidence of internal consistency for data from both versions of the questionnaire.

As displayed in Table 1, the ICC(1) values were computed for each item and ranged from .14 to .33, indicating substantial between-team differentiation (cf. Muthén & Satorra, 1995). In addition, all ICC(2) values were above .70 for the MATS-Global items as well as the MATS-Short items except for two of the adjustment items. The median $r_{WG}$ values across teams ranged from .60 to .80 for the MATS-Short and .68 to .80 for the MATS-Global. Overall, the items strongly differentiated between teams and produced reliable estimates of team-level scores.

**External Validity**

Multilevel correlations between teamwork and all correlates are shown in Table 2—the overall teamwork score provided is derived from the five-item MATS-Global, whereas the scores for each of the five aspects are derived from the five subscales within the 19-item MATS-Short (bivariate correlations and descriptive statistics are also available at the project’s OSF page: https://osf.io/pqnsw/). Regarding group constructs, our hypothesis of significant, large-sized correlations of teamwork with task cohesion (H2a) was supported in all cases except for athlete-level evaluation (which was medium sized). Our hypothesis of significant, medium-sized correlations with social cohesion (H2b) was supported in all cases except athlete-level evaluation (which was small in size) and team-level MTM (which was large in size). There was mixed evidence supporting our hypothesis of significant, large-sized correlations with collective efficacy (H2c)—large correlations were, indeed, found at the team level, but only medium-sized correlations were evident at the athlete level (although all correlations were significant). Finally, our hypothesis of significant, large correlations in relation to team performance satisfaction (H2d) was not supported. Specifically, at the athlete level, all correlations were significant but were either small (evaluation, adjustments, and MTM) or medium (general teamwork, preparation, and execution) in size. At the team level, only teamwork execution demonstrated a significant (medium-sized) correlation with team performance satisfaction; the remaining correlations were nonsignificant. The HTMT tests provided evidence of discriminant validity between data from both the MATS-Short and MATS-Global models in relation to all group constructs. Specifically, all HTMT values were within the 0.85 cutoff (Henseler et al., 2015), ranging from 0.21 (evaluation factor—team performance satisfaction) to 0.81 (general teamwork factor—task cohesion)—see the project’s OSF page (https://osf.io/pqnsw/) for a detailed summary.

Regarding individual constructs, teamwork had significant, small-sized correlations with individual performance satisfaction at the athlete level (thereby supporting H2e). At the team level, these correlations were significant and large. Mixed support was shown for the hypothesis of significant, small correlations with enjoyment (H2f). Specifically, there were, indeed, significant correlations at the athlete level, but aside from evaluation, these were medium in size. There were significant correlations at the team level for general teamwork, execution, adjustments, and MTM, but these were all large in size; in contrast, there were small and nonsignificant correlations for preparation and evaluation. Finally, our hypothesis of a significant, small relationship with commitment (H2g) was partially supported. All correlations were significant. At the athlete level, correlations were small in size for evaluation, adjustments, and MTM; they were medium in size for general teamwork, preparation, and execution. At the team level, all correlations were large.

We found partial support regarding our hypothesis of significant, medium-sized correlations between teamwork and transformational leadership (H3). All correlations were, indeed, significant. At the athlete level, these correlations were medium for general teamwork, preparation, execution, and adjustments and small for evaluation and MTM. A similar difference across these same factors was noted at the team level—the correlations were medium in relation to evaluation and MTM but large in relation to general teamwork, preparation, execution, and adjustments.

**Discussion**

The purpose of this study was to examine the structural and external validity of data derived from two shortened versions of the MATS, which assesses athletes’ perceptions of teamwork within their sport team. There was good model–data fit overall as well as evidence of internal consistency for both hypothesized models we tested—one relating to a 19-item version (i.e., the MATS-Short) and one relating to a five-item version (i.e., the MATS-Global). Moreover, the questionnaires appear to be well suited to detect between-team differences and produce reliable team-level estimates, as supported by the ICC(1), ICC(2), and $r_{WG}$ statistics. For the most part, there was evidence of external validity insofar as data from those models were positively and significantly related to (yet discriminant from) data assessing various group (team cohesion and collective efficacy) and individual (commitment and individual performance satisfaction) constructs. Although the size of some of these associations at times differed from our a priori hypotheses, they were all directionally consistent, and only a few (team-level) correlations—namely, in relation to team performance satisfaction and enjoyment—were not significant. Finally, significant positive correlations between teamwork and coach transformational leadership were evident, which is a novel finding to the teamwork in sport literature.

**Implications for Measuring Teamwork in Sport**

In light of the evidence supporting model–data fit and internal consistency in this study and previous research (McEwan et al., 2018), it would appear that researchers could use the 66-item...
MATS, 19-item MATS-Short, or 5-item MATS-Global, depending on their research question(s). A notable strength of the MATS is that it can provide estimates of one or more of the 14 dimensions of teamwork and/or one of more of the five aspects of teamwork (each comprising multiple dimensions). Hence, those seeking to examine teamwork at the dimension level should do so with the original version of the MATS rather than either of the shorter versions. Alternatively, the MATS-Short may be ideal for those seeking to assess one or more of the five aspects of teamwork in a more expedient way. For example, researchers may be keen to measure teamwork alongside multiple other constructs at the same time point(s). In such cases, the 19-item MATS-Short may be better suited than the 66-item MATS. Although the 19-item version would not provide the same level of detail as the 66-item version, it would allow teamwork to be measured at the same time as other variables without overburdening participants.

In addition to the 66-item MATS and 19-item MATS-Short, the five-item MATS-Global offers a third option for measuring teamwork. Although this measure can only offer a single estimate based on participants’ ratings of their team’s overall teamwork preparation, execution, evaluation, adjustments, and MTM, obtaining an omnibus score (or “snapshot”) of teamwork in this manner could provide opportunities for researchers to investigate the construct in new and creative ways. As one example, it is conceivable that this five-item version could be adopted to analyze the structure of relationships within a team in terms of its teamwork (e.g., through social network analysis) or interpersonal sources of teamwork perceptions (e.g., via social relations modeling). As another example, ecological momentary assessment (EMA) has recently gained traction within sport and exercise psychology (see Reifsteck et al., 2021). Considering that teamwork is conceptualized as a dynamic process (McEwan & Beauchamp, 2014), an EMA study that gauges athletes’ perceptions of teamwork behaviors repeatedly (e.g., on each day of a team’s practice or competition) would provide much greater insight into the episodic cycles of, and developmental influences on, teamwork beyond typical cross-sectional designs. Conducting EMA with the 66-item MATS or 19-item MATS-Short would likely be unfeasible due to participant burden; by measuring teamwork with five items, researchers are now much better positioned to pursue such endeavors. In summary, when deciding which of the three versions of the MATS to use, researchers will need to balance the nature of the research question(s) being posed with feasibility and the level of detail they seek to obtain in their teamwork assessment.

**Correlates of Teamwork**

Aside from a few exceptions (discussed later), there was general support for our a priori hypotheses in terms of statistically significant relationships between teamwork and its correlates. However, the findings were more mixed in terms of the effect sizes of those correlations. As a general note, we observed that the majority of team-level correlations were larger than their individual-level counterparts for each variable. With regard to group constructs, we found that teamwork generally had significant, large correlations with task cohesion and significant, medium correlations with social cohesion. These findings supported our a priori hypotheses and aligned with previous research (McEwan, 2020; López-Gajardo, Díaz-García, et al., 2023). Hence, we continue to gain more evidence of a large relationship of teamwork and unity around a team’s task objectives. What appears less clear, however, is the relationship between teamwork and another emergent state: collective efficacy. Specifically, although our findings at the team level align with our hypotheses and with previous research (McEwan, 2020; López-Gajardo, Díaz-García et al., 2023) insofar as teamwork demonstrated large, significant correlations with collective efficacy, we only found medium (though still significant) correlations at the individual level. Even less clear is the relationship with team performance satisfaction; although we anticipated large, significant relationships between teamwork with this output, we only found small- or medium-sized correlations at the athlete level. Moreover, only teamwork execution had a significant association (to a medium extent) with team performance at the team level.

Although further work is needed, we can speculate on potential explanations regarding the differences in size between team- and individual-level correlations as well as the results concerning the associations of teamwork with collective efficacy and team performance satisfaction. First, the differential findings in our study compared with previous work (McEwan, 2020) may be due, in part, to differences in the data analytic strategies between these studies. Whereas McEwan (2020) only examined individual-level associations between teamwork and external variables, we calculated multi-level estimates that partitioned variance at the team level (i.e., team scores) from the athlete level (i.e., latent scores reflecting deviations from the team score). In other words, the team-level correlations in our study are based on latent group means (e.g., the extent to which a team’s aggregate teamwork score relates to its mean collective efficacy score), whereas the athlete-level estimates represent individual deviations within teams. Given that teamwork is a group construct and the ICC values for teamwork and the correlates were substantive, one might reasonably expect that correlations and variability would be larger at the between- than within-team level. From a conceptual standpoint, another possibility—albeit speculative—is that athletes might vary in how they define success within a given team, which could weaken team-level associations. For instance, some players might focus on how well the team performed relative to its potential and previous efforts whereas others focus more on the team’s performance relative to opponents (e.g., win–loss record and placement in league standings).

Second, it should be reiterated that in the research by McEwan (2020), teamwork was measured at a different time point than collective efficacy and team performance satisfaction, with the latter two assessed closer to the end of the participating teams’ seasons. In contrast, the current study had a single time point, and teams were able to participate at any point in their season following the first month and at least two competitions. As such, in the case of team performance satisfaction in particular, it may have been premature in this study to ask participants to rate their satisfaction with their team’s success, overall performance, and degree to which they reached their performance goals for the season. Regardless, there appears to be accruing evidence that teamwork has a medium, rather than a large, association with team performance indices in light of the results from this study and others, such as Fransen et al. (2020) and López-Gajardo, Díaz-García, et al. (2023), who also found medium-sized correlations between these variables (at the athlete level). It also appears that teamwork execution is a prominent correlate of team performance whether measured in terms of satisfaction (e.g., the current study; Fransen et al., 2020; López-Gajardo, Leo, et al., 2023) or with more objective estimates (e.g., LeCouteur & Feo, 2011). This is perhaps unsurprising as the execution phase occurs during gameplay (i.e., “action” episodes) and, therefore, has the most proximal influence on team performance (Marks et al., 2001). In contrast, the relationships between the other four aspects...
of teamwork with team performance outcomes seem less certain and, therefore, require further research attention.

With regard to individual constructs, teamwork was consistently associated with athletes’ satisfaction with their personal performance and their commitment to their team, which aligns with previous research (McEwan, 2020). At the within-team level, these correlations were small for individual performance satisfaction and small or medium for commitment; at the between-team level, all correlations were large. As noted in the previous paragraph, the differences in the sizes of those individual- and team-level relationships may be partly due to teamwork being a group construct and the differences between teams being more prominent than the variability within teams. Significant correlations between teamwork and enjoyment were also evident at the individual level. However, at the team level, the correlations were large and significant for general teamwork, execution, adjustments, and MTM but small and nonsignificant for preparation and evaluation. There do not appear to be any conceptual or evidence-based reasons for these differences in statistical significance, and caution is warranted in interpreting the findings due to the cross-sectional nature of this study. Nonetheless, it might hint that further study should be devoted to disentangling these relationships and assessing whether execution, adjustments, and MTM are more salient predictors of athlete enjoyment compared with preparation and evaluation. These findings also further underscore the value and importance of conducting multilevel, rather than single-level, examinations of the relationships between teamwork and associated variables.

This study is the first to provide quantitative evidence of a relationship between coaches’ transformational leadership and teamwork behaviors in sport. The correlations were mostly medium at the athlete level and large at the team level. The significant association between these two variables complements recent research evidencing the influence that coaches have in fostering effective teamwork in sport (e.g., Fransen et al., 2020; Sokoloff et al., 2023) and also aligns with findings from industrial/organizational psychology regarding transformational leadership specifically (Burke et al., 2006). That said, an important caveat is that all of the work to date on coach leadership and teamwork in sport is based on data obtained at a single time point; as such, these studies do not capture temporal influences or dynamic associations between these variables (e.g., the extent to which coach leadership prospectively predicts teamwork; the ways in which this relationship changes over time). Hence, although our results add to the initial insight into the apparent influence that coaches have on teamwork, further investigation is needed to provide more definitive conclusions on this relationship.

Limitations and Additional Future Research Directions

Despite the contributions of this study to the teamwork in sport literature, there are important limitations that should be considered. For one, it is recognized that the items for the MATS-Global are nested within the MATS-Short and that the current data used to assess the MATS-Global were captured in this manner (i.e., with the 19 items from the MATS-Short). As such, it would be beneficial to conduct further validation work specifically on the five-item global version of the MATS. In addition, as all variables were measured at the same time using self-report questionnaires, common method variance (Podsakoff et al., 2003) may have influenced our findings. Thus, now that support for the validity of data derived from the MATS-Short and MATS-Global has been obtained, researchers should turn to study designs that allow for a greater understanding of temporal influences and causal effects between teamwork and inputs, emergent states, and outcomes (e.g., longitudinal research, intervention designs, and EMA approaches). It should also be noted that the majority of participants competed at the university level. Although we sought to obtain as many professional- and national-level teams as possible, accessing these populations can be challenging (and proved to be so in this study). Nonetheless, efforts should continue to be made in future research to sample elite teams and assess whether our accumulating knowledge of teamwork generalizes to these high levels of competition. Relatedly, more diverse samples in future research would also contribute to the generalizability of this research area. This could include recruiting large samples from several countries to conduct cross-cultural measurement invariance testing on data derived from these measures. In addition, although the selection of correlates was guided by previous validation work in sport (i.e., McEwan, 2020) as well as meta-analytic findings from other contexts (i.e., LePine et al., 2008), it is acknowledged that a range of other variables could have been examined. Hence, a key opportunity for researchers in future remains to examine the link between teamwork and other variables that influence team effectiveness (e.g., social identity and collective coping). In doing so, the field will gain more comprehensive insight into the nomological network of teamwork.

Conclusions

In summary, our study offers two new, shorter ways of measuring teamwork in sport. The provision of these measures provides more flexibility for researchers seeking to examine this construct. Our hope is that these options in assessment will serve as a catalyst for exciting new endeavors in research and applied practice, which will continue to enhance our understanding of teamwork, its importance, and the influences on its development.

Acknowledgment

The research conducted and described in this manuscript was not supported by a funding agency.

References


(Ahead of Print)


