rijksuniversiteit groningen

What makes a team work? Psychometric properties of the Teamwork Checklist

Stijn de Zeeuw

Master Thesis – Talent development and creativity

S3424316

June 2023

Department of Psychology

University of Groningen

Examiner/Daily supervisor:

Mark Schuls

Contents

Abstract

The Teamwork checklist (TC) is a questionnaire made for evaluating team dynamics and has been used for over a decade. The TC consists of six subscales: Cohesion, Competent leadership, Coaching, Commitment, Collective Goals and Coping and is used for advising teams on how to improve effectiveness.

This study examined the Teamwork Checklist and evaluated its construct validity, internal consistency reliability and test-retest reliability. The hypothesis was that the TC is a valid and reliable mode of measurement for team dynamics.

The test-retest reliability was assessed by testing six korfball teams twice over a course of four weeks. And was tested by using Pearson's product moment correlation for individual results and Spearman's Rho for team mean results. Team mean reliability was inconclusive because of sample size issues. Individual test-retest reliability showed that Coaching, Competent leadership and Coping were test-retest reliable and the other three subscales were not. Construct validity was calculated using principal component analysis (PCA) with an oblique rotation. The subscales showed correlation between each other, but not enough to doubt discriminative validity. The PCA showed clear difference between subscales, except for the subscale Commitment. Internal consistency reliability was acceptable for each subscale (at least α=0,75)(appendix 8). Additional alpha calculations showed room for removing at least one item from every subscale without sacrificing reliability (appendix 9).

Based on validity and reliability tests in this study, it can be concluded that the TC has construct validity and internal consistency. Test-retest reliability is less conclusive, one explanation for this could be a recency bias for social constructs. Future research could be aimed at establishing convergent validity and shortening the questionnaire to increase the ease of completing the TC.

Introduction

For a team to perform well on each level it is important for the individual players to cooperate well together. While this is not new information, the way to achieve this has been a topic of some debate. Historically, research has been scarce and inconclusive on the relationship between team dynamics and team performance in sport. Team dynamics and group performance has been studied in work and education contexts from the middle of past century, but Mullen et al. (1994) was the first study to generalize team dynamics towards a sport context. In this study the relationship between team cohesion and team performance was examined. However, the relationship between team cohesion and performance has never been confirmed, with multiple studies arguing for and against the effect. This debate did however spark the research into team dynamics as a function of performance. In the years that followed, the research of Mullen et al. (1994) multiple theories were formed on how team dynamics can influence performance. In the next section a brief literature study will be conducted in order to highlight the need for investigating team dynamics and the need for reliable and valid measures of team dynamics. The goal of this study is to relate this literature study to the Teamwork Checklist (TC) and to investigate the construct validity, test-retest reliability and the internal consistency of the TC.

After Mullen et al. (1994) generalized team cohesion from multiple different disciplines towards sports reactions were mixed. Because sport has a multidimensional nature, it might not be valid to generalize older research from other disciplines to a sport setting (Carron et al., 2002). Carron's study therefore attempted to establish a direct correlation between cohesiveness and performance and showed that sport teams are generally more cohesive than other disciplines (like military groups and groups which are used in organizations).

While it has been established that sport teams usually are highly cohesive, later research proved that a team does not have to be cohesive in order to be effective (McEwan & Beauchamp, 2014). This does not dispute the calls for new research into team dynamics, but

did show that new ways have to be explored in order to analyze the effectiveness of teams (Carron et al., 2012; McEwan & Beauchamp, 2014). One of the first calls for change was to change the broad term of teamwork to team effectiveness. This subtle change was important because "teamwork" implies a cohesive team and "effectiveness" is meant to shift towards the performance a team achieves, of which cohesion is only one dimension. Teamwork is part of team effectiveness, but only focuses on the behavior of the players (McEwan & Beauchamp, 2014). McEwan & Beauchamp state that team effectiveness can be seen as a linear model with input, mediators, outcomes and three dimensions of team effectiveness that influence performance. These dimensions are preparation, execution, and evaluation. Preparation regards out of field preparation before the main task of the team, meaning that goals are set for the next match and that team members are aware of these goals. Task execution can be divided into three facets and concerns in field behavior of players. The first facet McEwan & Beauchamp found was coordination, which concerns managing the independent and interdependent actions of team players in the field. Cooperation is the second facet and focuses on collective task execution. Coordination and cooperation differ from each other because cooperation implies that team members work together towards a common goal for mutual benefit where coordination requires team members to keep track of each other's actions. Cooperation also requires the team members to work together towards a certain goal. The last facet McEwan & Beauchamp found was communication which is the extent and manner of how team members share information with each other.

The final dimension McEwan & Beauchamp specified in team dynamics was evaluation and adjustment. This dimension concerns the monitoring of set goals and tracking progress towards team goal attainment. This is an out of field dimension and is not only about evaluation but also involves changing the process to make more progress towards the common team goal.

This dimension uses feedback, analysis and observing teammates in order to help each other improve.

An important part of the meta-analysis of McEwan & Beauchamp (2014) was that they considered team effectiveness as a process that happens in a vacuum within a team. In the last dimension it was stated that the model only takes into account that team members will evaluate peers in order to attain certain goals. In the study it is also stated that team dynamics take effect into broader external contexts, like supporting staff, but did not involve these factors into the model. Other theories on team dynamics have taken the angle of supporting staff into account, one of which was a study done by Janssen et al. (1999). Janssen studied team dynamics and the creation of a high-performance team. Based on this article Janssen also authored a popular psychology book (Janssen, 2002). In this book, Janssen states that a team which scores high on seven C's has a high probability of performing well. The seven C's stated are: Common goals, Commitment, Complementary roles, Communication, Constructive conflict, Cohesion, and Credible coaching (Janssen, 2002). While easily implementable, the scientific world did not really adopt Janssens' theory of the seven C's and tends to stick with McEwan & Beauchamp (2014).

The teamwork checklist (TC) is a questionnaire loosely based on the work of Janssen et al. (1999). The TC was designed by Mark Schuls, an applied sport psychologist, as a way to measure different assets that a team needs in order to operate effectively. The TC contains 36 items and was constructed by adding questions regarding coping/resilience to Janssens initial questionnaire and then performing an exploratory factor analysis in order to regroup the questions (Sarkar & Fletcher, 2013). From this analysis, Schuls defined six C's of team dynamics: Cohesion, Collective goals, Commitment, Coping, Competent leaders, and finally Coaching. These six C's will be defined and explored further in this study. The six C's of the TC fit within both Janssens study (1999), as the work of McEwan & Beauchamp (2014). The

TC combines the value of coaches which was described by Janssen with the dimensions of team effectiveness McEwan & Beauchamp found.

Cohesion as a concept is twofold. Firstly there is task cohesion, highlighting the amount as to which a team is able to work together towards common goals. The second type of cohesion is social cohesion, the extent as to which a members of the group are attracted to each other and to the goals they want to reach (Wolf et al., 2015). In terms of the 6C's, cohesion is regarded as social cohesion, being the social part of working together. Task cohesion will be reflected in other parts of the instrument. When a team is considered as being socially cohesive, they will not only work together well as a team, but also will like to meet outside of a sports context. Players in a cohesive team will accept others' personal differences. Previous research has indicated that there is a link between social cohesion and team performance (Carron et al., 2002; Mullen & Copper, 1994). The strength of this link has been disputed in recent years, but recent research has investigated new structures which support the investigation of social cohesion when finding points of improvement within teams (McEwan & Beauchamp, 2014; Zepp & Kleinert, 2015).

One of these structures is the presence of homogeneity of prototypical attributes within teams (Zepp & Kleinert, 2015). Prototypical attributes, like perceptions, behaviors, norms, rules, etc. are consciously or unconsciously synchronized in a cohesive team. When these attributes are heterogenous, more group members will experience a level of misfit and teamwork usually deteriorates (Harrison & Sin, 2007). Size can be a moderator for this because bigger teams are more likely to have homogeneity of attributes. However, while a cohesive team is mostly viewed as positive, too much cohesiveness will result in time wasting or communication issues, such as avoiding conflict to maintain cohesiveness (Hardy et al., 2005). Moreover, competitiveness can strain cohesiveness. It is therefore important to monitor the

cohesiveness of sports teams in order to walk a fine line between too cohesive and not cohesive enough.

"Collective goals" refers to the amount of communication a team has had on the goals the team is setting. Collective goals are usually labeled under how task cohesive a team is, together with commitment, which will be talked about more in detail later in this article. The use of goal setting, in particular difficult and specific goals, was first described by Locke and Latham (1984). This was first just based on the individual but more recently also the setting of team goals has also been studied (Senécal et al., 2008). Goal setting is frequently used by athletes in order to be able to measure progress, and to stimulate higher performance (Forsblom et al., 2019). For both individuals and teams, goal setting is most useful when the athletes who are benefitting from the goals are able to measure their progress towards the goal.

Goals setting in teams is used to increase motivation and team commitment, because players pursue shared objectives. Goal setting therefore works best when the team is involved in deciding which goals are set and which priority each goal gets (Durdubas & Koruc, 2023; Forsblom et al., 2019; Locke et al., 1988). Teams tend to set multiple goals to which they agreed upon together and usually set outcome goals and process goals for short term (the training session) or long term for competition settings. Senécal et al. (2008) found that most teams who use goal setting (38 out of 41 investigated) experience a positive performance enhancement from goal setting compared to teams who do not use (effective) goal setting. The other benefit from goal setting is that it helps increasing and maintaining a higher level of cohesion within the team. This could also be because most team building interventions contain some form of goal setting (Durdubas & Koruc, 2023). This means that there could be a cross-over effect from team building (Senécal et al., 2008).

Task cohesion is further represented in the TC under the commitment the players have towards achieving a certain goal. Players show this commitment by showing up for training and

matches. A team with strong goal commitment will consist of team members who agree to the goals that are set and have a shared determination to reach these goals (Han et al., 2018). Team goal commitment positively influences team performance and team viability (Aube & Rousseau, 2011). Commitment to the goals is also able to mediate the effects of in-team conflict and interpersonal aggression. Therefore, high goal commitment is able to act as a buffer to in-team conflict and therefore a high commitment to the goals is a preferable trait within teams. Commitment only works well when the members of the team are actively involved in the setting of the goals and if the goals that are being set consider the entire team. when players set goals for each other, but these goals are not team goals, the commitment towards these goals is less strong and effective in terms of its relationship with the performance.

Coping is defined as changing one's cognitive and behavioral efforts to manage specific demands that are appraised as taxing or exceeding the resources of the person (Leprince et al., 2018). The first theories on coping were about the transactional nature between stressors and the environment (Lazarus, 1999). In the concept of teams, coping is defined as the ability of a team to cope with setbacks and the mental flexibility of a team. In team sports, team members make use of communal coping (Leprince et al., 2019). Communal coping is defined as appraising and dealing with stressors in close relationships. Teams who have communal coping skills are able to deal with potential stressors collectively (Leprince et al., 2019) Most research into coping has been done for individual players, but some communal coping strategies have been identified. Recent research found various categories of stressors in a regular match, most of which had to do with interpersonal interactions or social encounters during a match (Holt & Hogg, 2002; Leprince et al., 2018). Research into basketball found that within teams, the stressors experienced during a game were shared, further building on the theory of communal coping. Because most team members share experienced stressors, Leprince et al. (2018) theorized that the coping of these stressors could also be handled in a group context. A

communal resilient team has been proven to be a key factor in performing under pressure and dealing with setbacks in matches (Morgan et al., 2013). Ways for teams to deal with stressors collectively involve sharing information about the stressor, building an action plan to deal with it, or increasing and restoring each other's confidence (Leprince et al., 2018). Apart from this, previous research identified a coping strategy which was unique to communal coping, being relationship-focused coping (Leprince et al., 2018). This way of coping involves maintaining the social relationship of the team during stressful moments, and maintaining the quality of intrateam relationships. Especially in research on how teams are able to perform at high levels and in stressful games, it is important to explore and enhance their individual and communal coping skills.

In the TC, Competent leadership investigates the faith and trust the team has in their leaders, both coaches and active players in the field (like team captains), and whether or not they experience team leadership as transformative or as servant leadership. Leadership is a complex and a multidimensional process, which leads to different leadership styes for different individuals. Teams who have high faith in their leaders usually have leaders who are able to adapt their leadership style to the needs and wants of their teams. Leadership can be roughly classified into three distinct forms of leadership, each with their own subclasses (Dijana et al., 2019). Autocratic leadership concerns leaders whose decisions are made without consulting their team members. Liberal leadership involves the team making decisions without the leader present. The type of leadership most apparent in sport teams is democratic leadership. This form of leadership involves leaders making decisions in cooperation with the team and allowing the members to take part in this process (Dijana et al., 2019). Team leaders can focus on the social side or the task based side of performance, good leaders improve the side they focus on. The TC measures the amount of experienced democratic leadership. Decisions are made in consultations with the team and the team chooses the leader themselves, to some degree.

This study will highlight two main theories in democratic team leadership: transformative and servant leadership (Price & Weiss, 2011; Worley et al., 2020). Transformative leaders are able to direct team members through behaviors in order to motivate them, whereas servant leadership focusses on fulfilling individual needs of team members (Worley et al., 2020). Previous studies have looked into leadership for both players or captains, and transformative leadership for supporting staff. Transformative leaders have been proven to lead to more intrinsically motivated players but servant leadership has a higher correlation with social cohesive teams (Worley et al., 2020). The other side of transformative leaders are laissez faire leaders. This style of leadership has been historically less effective and less motivating for sports teams.

The last C of the six C's in the TC is the level as to which the team members perceive the coaching of the supporting staff being beneficial to team results. Just as with leadership within teams, coaching is a multidimensional discipline and there are many different coaching styles. The TC investigates the extent to which the coach creates a task-involving environment (Harwood et al., 2015). Goal setting has been shown to improve motivation and improvement of players, especially when mastery approach goals are set. The coaching staff plays a vital role in creating the environment for setting mastery approach goals. According to achievement goal theory, there are two climates a coach can create for their team, task-involving and ego-involving (Monteiro et al., 2018). Task-involving emphasizes personal growth and encourages mistakes as a part of learning. This also means supporting players in setting mastery goals which have proven to be more effective. Ego-involved climates highlight competence and comparison between players. These climates also tend to punish mistakes. Ego-involving climates encourage performance goals and in these climates players tend to set avoidance goals, not achieving worse than their peers (Appleton et al., 2016). Research found that a task-involving climate allows for the development of more difficult tasks, and more motivated players. Earlier

studies also established that perception of coaching support was related to greater need fulfillment and more sport motivation (Jacqueline & Wilson, 2012).

The TC investigates the extent as to which the players feel the coach creates a task-involving climate and the style the coach uses for setting this climate. There are multiple leadership styles, but the TC investigates the referent power of the coach, whether or not the coach fits in the social identity of the team (Rylander, 2016).

Even though the TC has been used for almost a decade in practice, after a preliminary exploratory factor analysis in order to create the TC after which it was slightly changed, it has not yet been validated. The aim of this study is to validate this questionnaire and to investigate its reliability. This will be researched using both historical data and new research in order to investigate the TC. This study will investigate the test-retest reliability, the construct validity and the internal consistency of the TC. Based on the literature study, the TC is expected to have strong construct validity. Because the different subscales influence each other and the topics are closely related, we also expect high correlation between subscales. Furthermore internal consistency is expected to be high, because the after an exploratory analysis subscales were defined to be more concrete and topic related. Lastly, because test-retest reliability for the TC has not yet been investigated, a hypothesis can not yet be formed. There are expectations that some items will be less test-retest reliable, because teams grow more cohesive over a year, and coaches improve during the year.

Study one: Test-retest reliability

Participants

This study was conducted using six korfball teams consisting of 8 to 12 players for a total of 68 approached players (28 male, 40 female). Players responded anonymous and only the IP address from where the players responded was collected. In the first round 52 players

responded and in the second round 49, 31 players could be definitively matched between sessions and were included (Appendix 1). All teams were recruited from the local student korfball club the researcher belonged to. Players were included if they were assigned to a team, according to the administration of the korfball club. This meant that (longtime) injured players were also included in in this research. Korfball is a sport of Dutch origin in which males and females play in the same team. This has the benefit of being able to test both male and female players without having to add more teams. Teams that were included all played on an amateur level, the highest playing team played on the fourth level and the lowest playing team played on the eight level of The Netherlands. The final sample consisted of six teams with players aged between 18-30 (μ=22.7, SD=2.4), studying at either the University of Groningen or other universities or colleges located in Groningen. Participants were not compensated as the filling in of the test did not take enough time. However, as an extra incentive the trainer of the team could request a summary of the results with an explanations.

Materials

For this study, only the TC was used as a questionnaire. The TC consists of 36 questions about the six C's. Responses are given on a five point Likert scale (Completely disagree-completely agree). The TC also requests some personal information of the participants, age, level of play and hours trained in a week. This was the only material used and was also used for study two to four and will not be further discussed as research instruments. The TC was elaborated on in the introduction, all items of the TC are added in appendix 10.

Procedure

To test the test-retest reliability the TC was assessed on two separate occasions. The teams were approached directly before or directly after a training session and asked to fill in the questionnaire without discussing their answers with other players. They received a brief explanation about the study and on the contents of the questionnaire before filling it in.

Participants also were able to read study information (appendix 12) and signed an informed consent (appendix 13). Between the first and second session of responding, a waiting period of between four and six weeks was established in order to ensure the participants were not able to recollect their answers from the first time of responding to the questionnaire, but the time was not too long so that significant changes to the respondents could happen which would influence the test-retest reliability (Kurpius & Robinson, 2006). Because the researchers hypothesized that current standing in the competition could be a moderator for certain answers on the TC, the window of both answer sessions was set at a moment in the season where all teams experienced down time (Benson et al., 2016). To further ensure that competition standing would not influence the answers on the questionnaire, teams that were recruited for this study were all in different standing in their respective competitions, some were in the bottom whereas some were in the top of their competition. Teams were not selected on their standing in the competition. For data collection Qualtrics XM with a University of Groningen license was used (Qualtrics, Provo, UT, USA,).

Analysis

For data analysis SPSS version 28 was used. The test-retest reliability was calculated by calculating the correlation for the six C's separately on both team and individual level between both sessions of filling in the TC. For team level Spearman's Rho was calculated because the N would be quite small. The test-retest reliability on an individual level was calculated using Pearson product-moment correlation coefficient (r). Spearman's Rho assumes a measurement of at least interval level and interrelated variables. It is also more useful for calculating correlation with a lower N. Furthermore, descriptive statistics were calculated. SPSS was also used for the analysis in study two to four. Guidelines for the strength of the correlation before establishing test-retest reliability have been a topic of some debate over the last years but this study used the guidelines of Matheson and Granville (2019). This article reiterates the study of

Portney and Watkins (2015) that states that an r between 0.5 and 0.75 as "poor to moderate, 0.75 to 0.9 as "good", and above 0.9 as "acceptable. Along with this, the Cohen's D has to be at least 0.8.

Results

The first round of data collection yielded 52 responses, divided over six teams. The second round of data collection took place four weeks later and yielded 49 responses divided over six teams. Appendix 2 contains response data per team in session one and two. The most important assumption for the test-retest reliability is that the true score is measured over a short interval, which was adhered to.

Team scores were averaged on each of the six subscales (Appendix 2). The average of the response on the first session was then correlated to the average response per team on the second session. Results of the Spearman's Rho correlation are shown in appendix 2. Based on team averages, the correlation between Collective goals (ρ=0.886, p=0.02), Commitment (ρ=0.829, p=0.042) and Cohesion (ρ=0.943, p<0.005) can be classed as acceptable. Based on this study, no correlation between the subscales Cohesion, Competent Leadership and Coping was found on a team level.

In order to collect more data about the test-retest reliability, this study attempted to include individual data on the TC. Since this was not originally included in the questionnaire the researchers reconstructed individual respondents out of IP addresses, age, gender and teams. Out of 49 possible matches, this connection was made for 62 responses, resulting in 31 individuals. For 44 responses it was not possible to make a conclusive connection. Data for this analysis is shown in table 4. Coaching (r=0.825, p<0.001) and Competent Leadership (r =0.707, p<0.001) classify as good, Collective goals (r = 0.581, p=0.001) and Coping (r = 0.631, p<0.001) as moderate and Cohesion (r=0.391, p=0.029) and Commitment (r=0.383, p=0.033) as poor.

Study two: Construct validity

Participants

For this study, data from the first round of study one was used together with historical data from 2014 until 2023. In total, 816 participants were included in this study with a mean age of 21.78 (SD = 6.67). In total eleven sports were defined and tested and taken into account. Nine respondents did not define their sport, but did state that it was a team sport (appendix 11). In order to have a sample size with a relevant reliability, the lowest amount of complete responses was 385 (Smith, 2020).

Procedure

Data not collected in study one was collected by Drs. M. Schuls through his sports psychology agency. Teams requested an analysis of their structures and for that purpose filled in this questionnaire. Participants then gave consent for their data to be saved and used for research purposes. Apart from the analysis resources discussed in study one, Google forms was also used as a means of data collection. This data collection tool was used for all historical data which was not collected in study one. This will also be the case for study three and four

Analysis

When subscales show low correlation between scales and correct factor loadings the argument for good construct validity can be made. Based on theoretical arguments it was expected that subscales had a decent amount of correlation between them, but before conducting a factor analysis correlations between subscales were calculated in order to check which factor analysis should be conducted (appendix 4) (Carron et al., 2012; McEwan & Beauchamp, 2014; Mullen & Copper, 1994). Based on existing literature, correlations of 0.25 or lower were disregarded, 0.26-0.5 were seen as weak correlations, 0.51-0.75 are fair correlations and everything above 0.75 was treated as a high correlation when alfa was <0.05, this standard was used both for positive and negative correlations (Berg & Latin, 2008).

A factor analysis was used to find structures or subscales within a test. In this study an explorative factor analysis was conducted, which is used to find out if the data reflects the six C's of the TC. This factor analysis was performed to explore the amount of covariance between subscales of the TC. For the explorative analysis a principal component analysis (PCA) was performed. Because it would make sense that some of the results on the six C's correlate with each other, an oblique rotation was added (Chapman, 2018).

Results

Firstly, correlation between scales was calculated in order to asses the amount of intertest correlation was enough to constitute a PCA analysis (appendix 4). All inter-scale correlations were higher than 0.55, which is moderate to high. The subscale Commitment correlated higher with all other subscales, but highest with Collective goals ($r=0.680$, $p<0,001$). This did mean that the first assumption for a PCA (high intertest correlation) was met. Because a PCA is not robust against outliers, a command in SPSS was given to identify possible outliers. This was not the case, so this assumption was also met. Lastly, PCA assumes no missing values. Since there were no outliers, this assumption was also met. Before the principle component analysis was conducted, the data was subjected to a Kayser-Meyer-Olkin (KMO) and a Bartlett test of sphericity in order to assess the suitability of the data for factor analysis. The KMO test showed a sampling adequacy with a high relationship among variables (KMO = 0.941). The Bartlett test investigates if correlation between items was significant ($\chi^2(630) = 12682.6$ ($p<0.001$)). A scree plot was also generated in order to visualize the factor loadings (Appendix 5). From this can be concluded that the data was, as expected, suitable for factor analysis. Because of the interfactor correlation, an oblique (Oblimin) rotation was performed.

After the subscale correlation a factor analysis was conducted. Appendix 6 shows the obtained pattern matrix. Only items with factor loadings higher than 0.30 are shown. Factor one contains the questions on subscale Collective goals and two questions of Commitment. This

scale was very robust with and Eigenvalue of 11.14 and accounted for 30.93% of variance. Factor two contains all questions about the coaching of the staff. This second factor also was robust with an Eigenvalue of 2.74 and 7.61% of explained variance. Factor three contained all questions related to the team cohesion. This factor has an Eigenvalue of 2.03 and an explained variance of 5.63%. Factor four contained questions for the subscale Coping and one question for the subscale collective goals, however this question had a bigger loading on factor one. This factor had an Eigenvalue of 1.918 and explained 5.63% of all variance. Factor five contained all questions on the subscale competent leaders and had an Eigenvalue of 1.417 with 5.33% of all variance explained. Finally, factor six contained four items of commitment. This factor had an Eigenvalue of 1.012 with 2.81% of variance explained. Factor 7 was not included because the Eigenvalue was lower than one. The questions of subscale commitment which loaded on the subscale common goals asked about the priority of the success of the team and amount of work the team was prepared to do in order to win when necessary

Study three: Internal consistency reliability

Analysis

The final study conducted was the analysis of the internal consistency reliability. This was measured using Cronbach's lower bound to reliability α. This has been proven to be a reliable measure of internal consistency, and is most widely used for measuring the internal consistency of a questionnaire. Based on COTAN guidelines, a reliability above 0.7 will be treated as "good", between 0.6 and 0.7 will be treated as sufficient, and everything below 0.6 will be set as insufficient for test-retest reliability (Egberink et al., 2023). Cronbach's alpha assumes equal modes of measurement between items, because all items are answered on a five point Likert scale, this assumption is met. The regular Cronbach's alpha was calculated for each subscale (the six C's). After this the consistency of the full TC was calculated using the stratified Cronbach's alpha (Cronbach et al., 1965). The stratified alpha uses the alpha of the subscales

to estimate the internal consistency of the full TC. This Alpha can not be calculated in SPSS so this was done by hand using formula 1 (Appendix 7). Finally, in order to investigate possible improvements to the TC, Cronbach's alpha was calculated for each subscale when one item in the subscale was removed.

Results

The internal consistency analysis showed that all subscales are reliable based on this test. All subscales had at least a Cronbach's Alpha of 0.7 (appendix 8). Coping had the highest reliability (α= 0.885)(appendix 8). however when taking into account COTAN guidelines, all subscales can be considered as adequate. Table 4 shows all alpha values and standard deviations for all subscales. After the reliability analysis of the subscale the stratified Cronbach's Alpha was calculated (Widhiarso, 2010). The stratified alpha of the TC was 0.948 which means that the entire TC is a reliable mode of measurement. Cronbach's alpha also stayed high when items in a subscale were removed (appendix 9). This means that certain items from each subscale could be removed without sacrificing internal consistency, resulting in a shorter questionnaire.

Discussion.

This study investigated the TC and conducted three psychometric tests in order to test its validity and reliability. The TC was evaluated on its test-retest reliability, construct validity and internal consistency reliability. This section will draw some conclusions out of the results and give some recommendations on direction for future research.

Test-retest reliability of the TC was investigated using six korfball teams of a local sports club. In order to assess test-retest reliability, COTAN guidelines were used, however there is some room for interpretation on these guidelines because they have been a topic of some debate over the last years (Matheson & Granville, 2019). Based on the results of the individual answers on the TC, that Coaching, Competent leadership and Coping can be classed as reliable over a period of four weeks. The other three subscales (Collective goals,

Commitment and Cohesion) can not be classed as reliable, based on this study. Test-retest reliability on a team level can not be assessed based on this study alone. three subtests, Collective goals, Cohesion, and Commitment were showed significant correlation between the first and second session, but the confidence intervals were big enough that a definitive conclusion cannot be drawn. The other three subscales did not correlate with between the first and second session.

Response on the questionnaires was at least 65% for both sessions, which is a good response. There could be multiple reasons for the low test-retest reliability of Collective goals, Commitment and Cohesion. These subscales can be summarized as task and social cohesion (Boyd et al., 2014). Previous research has shown that cohesion is heavily influenced by a recency bias, which means that the last meeting the team had informs the answers given on these subscales. This study was conducted during down-time in the season, but players were not kept separate during this period, so interactions outside of matches can also play a role, for example, general unhappiness within a team or past social events. This could be mitigated by instructing players to keep in mind that they are reviewing the season thus far, and not just the last interaction.

When the TC was investigated for construct validity, The PCA analysis showed that subscales of the TC are distinguishable based on internal structure. Subscales loaded on different factors, showing that the subscales of the TC fit well with the model of the PCA. Based on this the argument can be made that the construct validity of the TC is high, and that the subscales investigate different constructs of team effectiveness.

The PCA analysis did show that two items of the Commitment subscale loaded on the subscale Collective goals. Both of these items had a lower loading than 0.3 on the factor of Commitment. While this confounding of subscales is unfortunate, it is not surprising, based on the earlier literature study. Both the setting of collective goals and team commitment are aspects

of task cohesion (Graupensperger & Tisak, 2018). Task cohesion in most literature is taken as a heterogenous construct, but in order to identify issues with task cohesion, the TC split task cohesion into two subscales. These subscales are close to each other based on theoretical arguments, and this study shows once again that confounding these two aspects of task cohesion is a threat when measuring these constructs independently. In order to still be able to measure both of these constructs separately without introducing confounding variables, the questions which load on a different factor could be rephrased or removed.

The third study calculated Cronbach's lower bound in order to test the internal structure of the subscales, and the stratified alpha in order to calculated to assess internal consistency of the TC as a whole. The alpha of the subscales and the stratified alpha was high. All subscales scored "good" based on the guidelines of COTAN (Egberink et al., 2023). However, based on appendix 1, alpha is high enough that it is possible to remove some items of the TC without sacrificing the internal consistency of the subscales. Future research could look into removing one item for each subscale, in order to increase the ease of filling out the questionnaire and reducing drop-off of participants and to further increase construct validity.

It is important to express some reservations to this study. The first is about the historical data which was used for this research on the construct validity and the internal consistency of the TC. Most participants of this questionnaire filled in the questionnaire with the goal of group based feedback to the coach on certain topics of team dynamics. This could change the answer some participants would give versus when they would fill it in when results are not disclosed with third parties after filling in. However, because every questionnaire was filled in voluntary and anonymously and with the goal of improving the team, it can be assumed that participants filled in the questionnaire honestly. This limitation can also be seen as a strength of the TC, because if players know that they have complete anonymity could help with filling out the questionnaire honestly.

In collecting the data almost no personal information was collected, and the questionnaire was filled in using anonymous links. This was a limitation because it made matching individual data from the first to the second session more difficult. Using provided age lists and IP addresses, individual data could be collected for 31 of the participants. This proved enough to calculate the correlation and draw meaningful conclusions, but was not as high as was aimed for at the start of the research.

A third limitation of this research contains limited testing pool for the test-retest reliability. In the end six teams filled in the questionnaire to be analyzed. Because scores are pooled per team, it is difficult to draw definite conclusions. This reliability could be further analyzed by adding teams which are tested multiple times a year. Because this questionnaire is only used to measure team average, it is good to investigate whether or not the TC is stable over a short amount of time. Furthermore, participants expressed a reduced motivation to fill in the questionnaire a second time because of the time it took to fill it in. This could lead to a selection bias. A way to deal with this limitation is to shorten the questionnaire in order to make filling in the questionnaire twice more approachable.

Even though this study has shown that the TC shows signs of being a valid and reliable measure of the dynamics within teams, it is always necessary to keep validating the questionnaire in order to generalize the TC and make the questionnaire more readily available. Future research could therefore look into examining the convergent validity of the TC. One of the questionnaires which can be used for this, is the MATS (McEwan et al., 2018). This questionnaire has been extensively validated and has also been shown to be a valid and reliable measure of team dynamics. However, the MATS is with 70 items quite large and can take a while to fill in. The TC could fill a gap by also measuring team dynamics but being a smaller questionnaire, it has a smaller barrier to filling in.

A final interesting future direction for research for the TC is by investigating how the players who filled in the TC behave while playing in the field. Because the TC is a self-report questionnaire, there is always a threat of self-presentation, meaning that players will fill in the TC based on how they would like to act instead of how they usually act in match situations (Niessen et al., 2017). This could be investigated by having teams fill in the TC and then testing for self report bias by systematically observing players of a team in an actual match. Investigating in-field behavior and knowing how players act versus what they fill in can also give insight in the evaluation dimension of team effectiveness (McEwan & Beauchamp, 2014).

Practical implications

The practical implications of this study are that some recommendations have been made in order to improve and further use the TC. The most important discovery made by this study is that two subscales of the TC confound each other, and it is recommended that the two questions which load on a different factor than they should are reframed or removed in order to improve the construct validity of the TC. Furthermore the practical recommendation is made to remove one item from each of the subscales in order to decrease the time it takes to fill in the TC and therefore increase the ease of use of this tool. When these recommendations are implemented, the next step in validating the TC is to investigate its actual effect in the field by conducting a study in which is investigating whether players act in the same way they filled in the TC.

References

Appleton, P. R., Ntoumanis, N., Quested, E., Viladrich, C., & Duda, J. L. (2016). Initial validation of the coach-created Empowering and Disempowering Motivational Climate Questionnaire (EDMCQ-C). *Psychology of Sport and Exercise, 22*, 53-65. https://10.1016/j.psychsport.2015.05.008

Aube, C., & Rousseau, V. (2011). Interpersonal aggression and team effectiveness: The mediating role of team goal commitment. *Journal of Occupational and Organizational Psychology, 84*, 565-580. https://https://doi-org.proxy-ub.rug.nl/10.1348/096317910X492568

Benson, A. J., Šiška, P., Eys, M., Priklerová, S., & Slepička, P. (2016). A prospective multilevel examination of the relationship between cohesion and team performance in elite youth sport. *Psychology of Sport & Exercise, 27*, 39-46. https://10.1016/j.psychsport.2016.07.009

Berg, K. E. 1., & Latin, R. W. 1. (2008). *Essentials of research methods in health, physical education, exercise science, and recreation* (3rd ed. ed.). Wolters Kluwer Health/Lippincott Williams & Wilkins.

Boyd, M., Kim, M., Ensari, N., & Yin, Z. (2014). Perceived motivational team climate in relation to task and social cohesion among male college athletes. *Journal of Applied Social Psychology, 44*(2), 115-123. https://10.1111/jasp.12210

Carron, Colman, M. M., Wheeler, J., & Stevens, D. (2002). Cohesion and Performance in Sport: A Meta Analysis. *Journal of Sport and Exercise Psychology, 24*(2), 168-188. https://10.1123/jsep.24.2.168

Carron, Martin, L. J., & Loughead, T. M. (2012). Teamwork and performance. (pp. 309-327). Oxford University Press. https://10.1093/oxfordhb/9780199731763.013.0016

Chapman, S. J. (2018). Review of Discovering Statistics Using IBM SPSS Statistics, 4th Edition. *Journal of Political Science Education, 14*(1), 145-147. https://10.1080/15512169.2017.1366328

Chen, S., Qin, P., Tan-Soo, J. -., & Wei, C. (2019). Recency and projection biases in air quality valuation by Chinese residents. *Science of the Total Environment, 648*, 618-630. https://https://doi.org/10.1016/j.scitotenv.2018.08.153

Cohen, S. G., & Bailey, D. E. (1997). What Makes Teams Work: Group Effectiveness Research from the Shop Floor to the Executive Suite. *Journal of Management, 23*(3), 239-290. https://10.1177/014920639702300303

Cronbach, L. J., Schönemann, P., & McKie, D. (1965). Alpha coefficients for Stratified-Parallel Tests. *Educational and Psychological Measurement, 25*, 291-312. https://10.1177/001316446502500201

Dijana, M. M., Simovic, O., & Raicevic, M. (2019). The Relationship between Leadership Styles and Organizational Culture in Sport Organizations. *Sport Mont, 17*, 89. https://10.26773/smj.191002

Durdubas, D., & Koruc, Z. (2023). Effects of a multifaceted team goal-setting intervention for youth volleyball teams. *Journal of Applied Sport Psychology, 35*(2), 224-243. https://10.1080/10413200.2021.2021564

Egberink, I. J. L., Leng, W. E. d. & Vermeulen, C. S. M. (2023). *Cotan Documentation.* Retrieved 3-3-2023, from www.cotandocumentatie.nl

Fasold, F. (2021). Gaze Coordination of Groups in Dynamic Events - A Tool to Facilitate Analyses of Simultaneous Gazes Within a Team. *Frontiers in Psychology, 12*https://https://doi.org/10.3389/fpsyg.2021.656388

Forsblom, K., Konttinen, N., Weinberg, R., Matilainen, P., & Lintunen, T. (2019a). Perceived goal setting practices across a competitive season. *International Journal of Sports Science Coaching, 14*(6), 765-778. https://https://doi.org/10.1177/1747954119880994

Forsblom, K., Konttinen, N., Weinberg, R., Matilainen, P., & Lintunen, T. (2019b). Perceived goal setting practices across a competitive season. *International Journal of Sports Science & Coaching, 14*(6), 765-778. https://10.1177/1747954119880994

Graupensperger, S. A., & Tisak, M. S. (2018). Moral Behavior in the Locker Room Predicts Perceptions of Task Cohesion in Youth Ice Hockey Player. *Journal of Human Kinetics, 65*, 273-280. https://10.2478/hukin-2018-0047

Han, S. J., Lee, Y., Beyerlein, M., & Kolb, J. (2018). Shared leadership in teams: The role of coordination, goal commitment, and knowledge sharing on perceived team performance. *Team Performance Management, 24*(3-4), 150-168. https://10.1108/TPM-11-2016-0050

Hardy, J., Eys, m., & Carron, a. (2005). Exploring the Potential Disadvantages of High Cohesion in Sports Teams. *Small Group Research, 36*(2), 166-187. https://10.1177/1046496404266715

Harrison, D. A., & Sin, H. P. (2007). Heterogeneity as misfit. *Perspectives on Fit in Organizations,* , 372-376.

Harwood, C. G., Keegan, R. J., Smith, J. M. J., & Raine, A. S. (2015). A systematic review of the intrapersonal correlates of motivational climate perceptions in sport and physical activity. *Psychology of Sport and Exercise, 18*, 9-25. https://10.1016/j.psychsport.2014.11.005

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*(3), 251-271. https://10.1123/tsp.16.3.251

Jacqueline, P. P., & Wilson, P. M. (2012). Understanding Motivational Processes in University Rugby Players: A Preliminary Test of the Hierarchical Model of Intrinsic and Extrinsic Motivation at the Contextual Level. *International Journal of Sports Science & Coaching, 7*(1), 89-107. https://10.1260/1747-9541.7.1.89

Janssen, J. (2002). *A winning game plan for building a championship team* (first ed.). Quality books, inc.

Janssen, J. (1999). *Championship team building : what every coach needs to know to build a motivated, committed cohesive team*. Winning the Mental Game.

Kleinert, J. (2012a). Group dynamics in sports: An overview and recommendations on diagnostic and intervention. *Sport Psychologist, 26*(3), 412-434. https://https://doi.org/10.1123/tsp.26.3.412

Kleinert, J. (2012b). Group dynamics in sports: An overview and recommendations on diagnostic and intervention. *Sport Psychologist, 26*(3), 412-434. https://https://doi.org/10.1123/tsp.26.3.412

Kurpius, S., & Robinson, E. (2006). *Testing and measurement : a user-friendly guide*. Sage Publications.

Latham, G. P. (1984). Goal setting works. *Gifts Decorative Accessories, 85*

Lazarus, R. S. (1999). *Stress and emotion : a new synthesis / Richard S. Lazarus*. Springer Pub. Co.

Leprince, C., D'Arripe-Longueville, F., & Doron, J. (2018). Coping in Teams: Exploring Athletes' Communal Coping Strategies to Deal With Shared Stressors. *Frontiers in Psychology, 9* https://www.frontiersin.org/articles/10.3389/fpsyg.2018.01908

Leprince, C., d'Arripe-Longueville, F., Chanal, J., & Doron, J. (2019). Development and preliminary validation of the Communal Coping Strategies Inventory for Competitive Team Sports. *Psychology of Sport and Exercise, 45*, 101569. https://10.1016/j.psychsport.2019.101569

Locke, E. A. (2006). New Directions in Goal-Setting Theory. *Current Directions in Psychological Science, 15*(5), 265-268.

Locke, E. A., Latham, G. P., & Erez, M. (1988). The determinants of goal commitment. *The Academy of Management Review, 13*, 23-39. https://10.2307/258352

Matheson, & Granville, J. (2019). We need to talk about reliability: making better use of test-retest studies for study design and interpretation. *PeerJ, 7*, e6918. https://10.7717/peerj.6918

Mathieu, J. E. (2015). Modeling reciprocal team cohesion-performance relationships, as impacted by shared leadership and members' competence. *The Journal of Applied Psychology, 100*(3), 713-34. https://http://dx.doi.org.proxy-ub.rug.nl/10.1037/a0038898

McEwan D, Zumbo BD, Eys MA, Beauchamp MR. (2018). The Development and Psychometric Properties of the Multidimensional Assessment of Teamwork in Sport. *Journal of Sport Exercise Psychology, 40*(2), 60-72. https://https://doi.org/10.1123/jsep.2017-0193

McEwan, D., & Beauchamp, M. R. (2014). Teamwork in sport: a theoretical and integrative review. *International Review of Sport and Exercise Psychology, 7*(1), 229-250. https://10.1080/1750984X.2014.932423

McGrath, J. E. (2007). Temporal Matters in the Study of Work Groups in Organizations. *The Psychologist-Manager Journal, 10*(1), 3-12. https://10.1080/10887150709336609

Monteiro, D., Teixeira, D. S., Travassos, B., Pedro Duarte-Mendes, João Moutão, Sérgio Machado, & Luís Cid. (2018). Perceived Effort in Football Athletes: The Role of Achievement Goal Theory and Self-Determination Theory.*9*https://10.3389/fpsyg.2018.01575

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

Mullen, B., & Copper, C. (1994). The relation between group cohesiveness and performance: An integration. *Psychological Bulletin, 115*(2), 210-227. https://http://dx.doi.org.proxy-ub.rug.nl/10.1037/0033-2909.115.2.210

Nicklas, A. (2022). Gaze behavior in social interactions between beach volleyball playersAn exploratory approach. *Frontiers in Psychology, 13*https://https://doi.org/10.3389/fpsyg.2022.945389

Niessen, A. S., Meijer, R. R., & Tendeiro, J. N. (2017). Measuring non-cognitive predictors in high-stakes contexts: The effect of self-presentation on self-report instruments used in admission to higher education. *Personality and Individual Differences, 106*, 183-189. https://10.1016/j.paid.2016.11.014

Portney, L. G., & Watkins, M. P. (2015). *Foundations of clinical research : applications to practice* (3rd edition, [revised] ed.). Pearson/Prentice Hall.

Price, M. S., & Weiss, M. R. (2011). Peer Leadership in Sport: Relationships among Personal Characteristics, Leader Behaviors, and Team Outcomes. *Journal of Applied Sport Psychology, 23*(1), 49-64. https://10.1080/10413200.2010.520300

Qualtrics, Provo, UT, USA.Qualtrics [computer software]

Rylander, P. (2016). Coaches Bases of Power and Coaching Effectiveness in Team Sports. *International Sport Coaching Journal, 3*(2), 128-144. https://https://doi.org/10.1123/iscj.2015-0046

Sarkar, M., & Fletcher, D. (2013). How Should We Measure Psychological Resilience in Sport Performers? *Measurement in Physical Education and Exercise Science, 17*, 264-280. https://10.1080/1091367X.2013.805141

Senécal, J., Loughead, T. M., & Bloom, G. A. (2008). A Season-Long Team-Building Intervention: Examining the Effect of Team Goal Setting on Cohesion. *Journal of Sport and Exercise Psychology, 30*(2), 186-199. https://10.1123/jsep.30.2.186

Weiner, B. (1974). *Achievement motivation and attribution theory*. General Learning Press.

Widhiarso, W. (2010). Estimate Reliability Measurement for Multidimensional Scales. *SSRN Electronic Journal,* https://10.2139/ssrn.1597532

Wolf, S. A., Eys, M. A., Sadler, P., & Kleinert, J. (2015). Appraisal in a Team Context: Perceptions of Cohesion Predict Competition Importance and Prospects for Coping. *Journal of Sport and Exercise Psychology, 37*(5), 489-499. https://10.1123/jsep.2014-0276

Worley, J. T., Harenberg, S., & Vosloo, J. (2020). The relationship between peer servant leadership, social identity, and team cohesion in intercollegiate athletics. *Psychology of Sport Exercise, 49*

Zepp, C., & Kleinert, J. (2015). Homogeneity of Prototypical Attributes in Soccer Teams. *SAGE Open, 5*(3)https://https://doi.org/10.1177/2158244015602517

Zucchermaglio, C. (2011). Identity in Sport Teams. *Psychology, 02*(03), 202-209. https://10.4236/psych.2011.23031

Appendix 1: Response on both sessions of the TC

Appendix 1.1

Response to the first and second round of data collection of the Teamwork Checklist.

Team	N Round 1	N Round 2	Total amount of players
1	6	7	8
2	9	9	12
3	9	8	12
4	9	8	12
5	10	9	12
6	9	8	12
Total	52	49	68

Appendix 2: averages TC and correlation

Appendix 2.1
Mean scores of subscales of the teamwork checklist on the first and second session and the Spearman correlation between both sessions.

		Mean (SD)				95% Confidence Intervals	
	N	Session 1	Session 2	ρ	p	Lower	Upper
Collective Goals	6	22.92 (3.62)	22.14 (3.2)	0.26	0.62	-0.72	0.89
Coping	6	20.21 (3.75)	19.65 (4.46)	0.60	0.21	-0.44	0.95
Competent Leadership	6	20.52 (3.97)	20.02 (3.75)	0.83	0.04	0.02	0.98
Commitment	6	21.10 (5.2)	20.71 (4.64)	0.60	0.21	-0.44	0.95
Coaching	6	21.16 (5.82)	20.47 (5.97)	0.94	0.01	0.54	0.99
Cohesion	6	20.81 (4.5)	20.16 (4.61)	0.89	0.02	0.23	0.99

SD = standard deviation. N = amount of teams. ρ = Spearman's Rho correlation between the first and second session. p = significance. Cohesion, competent leadership and coaching are significantly related to each other.

Appendix 3: individual test-retest reliability for TC

Table 3.1
Correlation between first and second response on the TC.

Subscale	N	r	p
Collective Goals	31	0.581	0.001
Coping	31	0.631	<0.001
Competent Leaders	31	0.707	<0.001
Commitment	31	0.383	0.033
Coaching	31	0.825	<0.001
Cohesion	31	0.391	0.029

N = number of respondents who could be identified as having filled in the questionnaire twice. R = pearson product moment correlation. P = significance. significance <0.05 is considered as acceptable

Appendix 4: subscale correlation

Table 4.1
Correlations between subscales.

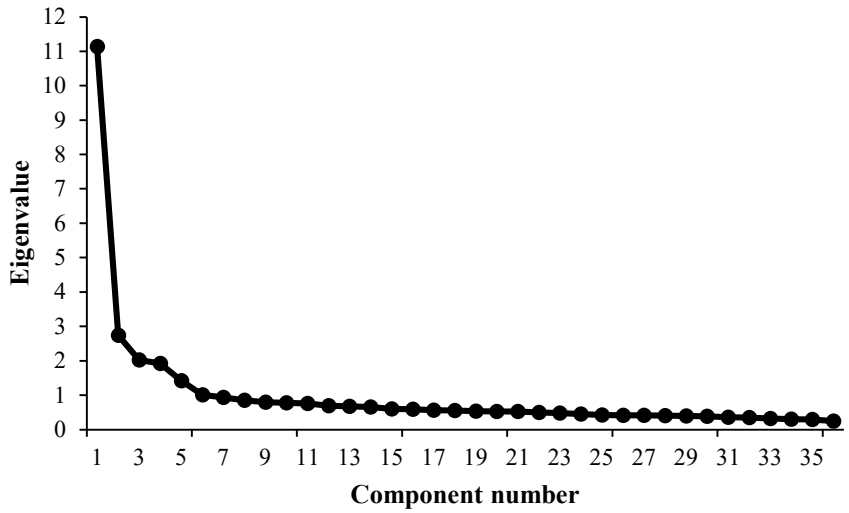
	Collective Goals	Coping	Competent Leaders	Commitment	Coaching	Cohesion
Collective Goals	1					
Coping	0.542	1				
Competent Leaders	0.464	0.427	1			
Commitment	0.680	0.514	0.446	1		
Coaching	0.459	0.370	0.595	0.416	1	
Cohesion	0.409	0.438	0.425	0.564	0.383	1

Correlations between 0.25-0.5 were seen as weak correlations, 0.51-.75 are fair correlations and everything above .75 is treated as a high correlation. All correlations were significant at p<0.05.

Appendix 5: Scree plot

Figure 5.1

Scree plot from Principal Component Analysis on the Teamwork Checklist



Components with an Eigenvalue lower than 1,0 were not selected, because this indicates that the factor is not robust enough.

Appendix 6: Factor loadings

Table 6.1
Principal component analysis on the TC (N=816).

Items	Factors					
	1	2	3	4	5	6
CGQ1	0.497					
CGQ2	0.806					
CGQ3	0.651					
CGQ4	0.815					
CGQ5	0.370					
CGQ6	0.403			-0,301		
CoachingQ1		0.736				
CoachingQ2		0.782				
CoachingQ3		0.553				
CoachingQ4		0.637				
CoachingQ5		0.574				
CoachingQ6		0.781				
CohesionQ1			0.543			
CohesionQ2			0.78			
CohesionQ3			0.621			
CohesionQ4			0.558			
CohesionQ5			0.502			
CohesionQ6			0.663			
CopingQ1				-0.859		
CopingQ2				-0.745		
CopingQ3				-0.781		
CopingQ4				-0.732		
CopingQ5				-0.818		
CopingQ6				-0.811		
CLQ1					-0.765	
CLQ2					-0.727	
CLQ3					-0.570	
CLQ4					-0.770	
CLQ5					-0.552	
CLQ6					-0.810	
CommitmentQ1						-0.579
CommitmentQ2	0.606					
CommitmentQ3						-0.622
CommitmentQ4						-0.570
CommitmentQ5	0.458					
CommitmentQ6						-0.615

CG = Collective goals. CL = competent leadership. Principal component analysis with oblique rotation. Loadings under 0.30 are not shown for readability.

Appendix 7: stratified alpha formula

Formula 7.1

Stratified Cronbach's lower bound to reliability.

$$\alpha_{Stratified} = \frac{\sum_{i=1}^{n} S_i^2 * (1 - \alpha_i)}{S_n^2}$$

N = number of tests, S^2= variance of subscale i. Sn^2= total variance of the test. αi = reliability of subscale i.

Appendix 8: Cronbach's alpha of the six subscales of the Teamwork Checklist

Table 8.1
Internal consistency on six subscales of the Teamwork checklist.

Subscale	Mean	Variance	Cronbach's Alpha
Collective goals	21.84 (3.72)	13.87	0.824
Coping	23.69 (3.91)	15.26	0.885
Competent Leaders	22.06 (3.83)	14.70	0.851
Commitment	21.69 (4.08)	16.68	0.787
Coaching	22.07 (4.10)	16.84	0.827
Cohesion	21.71 (3.64)	13.23	0.762

internal consistency was calculated using Cronbach's lower bound to reliability (alpha).
α > 0.75 is considered as adequate. Mean score (standard deviation). Each subscale consist of six items

Appendix 9: Reliability statistics with removed items

Table 9.1:
reliability statistics when each item is deleted for subscale Cohesion.

	Scale Mean if Item Deleted	Scale Variance if Item Deleted	Corrected Item-Total Correlation	Cronbach's Alpha if Item Deleted
CohesionQ1	18.31	9.54	0.45	0.743
CohesionQ2	18.50	8.81	0.52	0.724
CohesionQ3	18.15	9.12	0.58	0.704
CohesionQ4	17.89	9.67	0.59	0.707
CohesionQ5	17.81	10.57	0.39	0.754
CohesionQ6	17.89	10.02	0.50	0.728

Table 9.2
Reliability statistics when each item is deleted for subscale Competent Leadership.

	Scale Mean if Item Deleted	Scale Variance if Item Deleted	Corrected Item-Total Correlation	Cronbach's Alpha if Item Deleted
CLQ1	18.38	10.36	0.60	0.833
CLQ2	18.22	10.39	0.68	0.818
CLQ3	18.33	10.59	0.63	0.827
CLQ4	18.25	10.40	0.69	0.816
CLQ5	18.61	11.00	0.53	0.845
CLQ6	18.53	10.30	0.67	0.820

Table 9.3
Reliability statistics when each item is deleted for subscale Coping.

	Scale Mean if Item Deleted	Scale Variance if Item Deleted	Corrected Item-Total Correlation	Cronbach's Alpha if Item Deleted
CopingQ1	18.06	11.49	0.71	0.862
CopingQ2	17.95	12.15	0.68	0.868
CopingQ3	18.30	11.47	0.68	0.870
CopingQ4	18.07	12.11	0.67	0.870
CopingQ5	18.03	11.85	0.73	0.859
CopingQ6	18.04	12.01	0.70	0.864

Table 9.4
Reliability statistics when each item is deleted for subscale Commitment.

	Scale Mean if Item Deleted	Scale Variance if Item Deleted	Corrected Item-Total Correlation	Cronbach's Alpha if Item Deleted
CommitmentQ1	18.14	10.37	0.50	0.761
CommitmentQ2	18.06	10.11	0.51	0.760
CommitmentQ3	18.60	10.04	0.52	0.758
CommitmentQ4	18.43	8.99	0.61	0.735
CommitmentQ5	17.64	10.57	0.54	0.755
CommitmentQ6	18.34	10.14	0.53	0.756

Table 9.5
Reliability statistics when each item is deleted for subscale Collective Goals.

	Scale Mean if Item Deleted	Scale Variance if Item Deleted	Corrected Item-Total Correlation	Cronbach's Alpha if Item Deleted
CGQ1	19.69	11.08	0.58	0.799
CGQ2	19.65	10.36	0.59	0.798
CGQ3	19.76	10.72	0.61	0.791
CGQ4	19.61	10.68	0.68	0.777
CGQ5	19.86	11.59	0.56	0.803
CGQ6	19.87	11.39	0.55	0.805

Table 9.6
Reliability statistics when each item is deleted for subscale Coaching.

	Scale Mean if Item Deleted	Scale Variance if Item Deleted	Corrected Item-Total Correlation	Cronbach's Alpha if Item Deleted
CoachingQ1	18.23	12.74	0.58	0.803
CoachingQ2	18.51	11.46	0.65	0.788
CoachingQ3	18.46	12.35	0.56	0.808
CoachingQ4	18.48	11.75	0.66	0.786
CoachingQ5	18.25	12.76	0.56	0.807
CoachingQ6	18.43	11.55	0.36	0.803

Appendix 10: Teamwork checklist

1. De coaches creëren een positieve omgeving om onze doelen te kunnen behalen.
2. De spelers in het team nemen persoonlijke conflicten niet mee het veld op.
3. De spelers in dit team hebben hun rol geaccepteerd en zijn bereid deze uit te voeren.
4. Spelers van het team tonen respect voor elkaar door zelden te roddelen.
5. De coach communiceert eerlijk, helder en effectief.
6. Iedereen in het team werkt toe naar een gezamenlijk doel.
7. Het team heeft een duidelijk doel voor dit seizoen gesteld.
8. Het succes van het team heeft hoge prioriteit voor de spelers.
9. Ik heb er vertrouwen in dat het team ons doel kan bereiken.
10. De coaches en spelers luisteren goed naar elkaars mening.
11. De coaches zorgen ervoor dat iedereen het beste uit zijn/haar kwaliteiten haalt.
12. Het team heeft competente leiders. (met leiders worden de mensen uit de begeleidingsstaf, aanvoerder(s) en bepalende spelers bedoeld)
13. Ook als het tijdens wedstrijden tegen zit, blijft het team ervoor te gaan.
14. De leiders van het team passen goed bij de groep. (met leiders worden de mensen uit de begeleidingsstaf, aanvoerder(s) en bepalende spelers bedoeld)
15. Ik vind ons doel voor dit seizoen haalbaar en uitdagend.
16. Het team zet door bij de meeste tegenslagen die op ons pad komen.
17. Het team heeft er vertrouwen in dat de leiders het team goed begeleiden. (met leiders worden de mensen uit de begeleidingsstaf, aanvoerder(s) en bepalende spelers bedoeld)
18. Het team blijft bij tegenslagen in de wedstrijd vertrouwen houden in eigen spel.
19. De coaches stimuleren samenwerking tussen de spelers.
20. De afspraken die we hebben gemaakt helpen om ons doel te bereiken.

21. De coaches laten waardering blijken voor alle spelers binnen het team, ook voor de reservespelers.
22. De communicatie tussen de spelers onderling is eerlijk, helder en effectief.
23. De spelers in het team vertrouwen elkaar.
24. De spelers in het team zijn bereid om hun persoonlijke doelen op te offeren in het belang van het team.
25. Alle spelers geven 100% voor het team.
26. Het team blijft gemotiveerd na onverwachte teleurstellingen.
27. Het team is bereid om hard te werken.
28. De spelers in het team komen voor elkaar op als ze van buitenaf worden bekritiseerd.
29. Iedereen in het team is ervan overtuigd het doel van elke wedstrijd te kunnen bereiken.
30. De spelers in dit team hebben goed door welke rol zij moeten vervullen om het team succesvol te laten zijn.
31. Ik vind de leiders van het team geloofwaardig in hun rol. (met leiders worden de mensen uit de begeleidingsstaf, aanvoerder(s) en bepalende spelers bedoeld)
32. Het team blijft er na tegenvallers in geloven.
33. De leiders van het team kunnen goed aanvoelen wat er speelt in het team.(met leiders worden de mensen uit de begeleidingsstaf, aanvoerder(s) en bepalende spelers bedoeld)
34. Na een onverwachte tegenslag tijdens de wedstrijd, blijft het team gezamenlijk doorzetten.
35. Over het algemeen kunnen spelers elkaars verschillen accepteren.
36. De leiders van het team laten een daadkrachtig leidschap zien. (met leiders worden de mensen uit de begeleidingsstaf, aanvoerder(s) en bepalende spelers bedoeld)

Appendix 11: Investigated team sports

1. Floorball
2. Handball
3. Field Hockey
4. Baseball
5. Korfball
6. softbal
7. Football
8. Volleybal
9. Kaatsen
10. Basketball
11. Rowing

Appendix 12: Information about study
INFORMATIE OVER HET ONDERZOEK

VERSIE VOOR DEELNEMERS

"WHAT MAKES A TEAM WORK? INVESTIGATING THE VALIDITY AND RELIABILITY OF THE TEAMWORK CHECKLIST."

PSY-2223-S-0075

➢ Waarom krijg ik deze informatie?
U wordt uitgenodigd om mee te doen aan dit onderzoek naar de validiteit en de betrouwbaarheid van de teamwork checklist. Dit is een vragenlijst die kijkt naar de interne structuren binnen sport teams. U wordt uitgenodigd omdat u in competitie verband in een sport team speelt.

Dit onderzoek wordt uitgevoerd door:

- Stijn de Zeeuw, BSc. (master student)
- Mark Schuls, MSc. (Thesis begeleider en assessor)
- Dr. Ralf Cox (Tweede assessor).

➢ **Moet ik meedoen aan dit onderzoek?**

Meedoen aan het onderzoek is vrijwillig. Wel is uw toestemming nodig. Lees deze informatie daarom goed door. Stel alle vragen die u misschien heeft, bijvoorbeeld omdat u iets niet begrijpt. Pas daarna besluit u of u wilt meedoen. Als u besluit om niet mee te doen, hoeft u niet uit te leggen waarom, en zal dit geen negatieve gevolgen voor u hebben. Dit recht geldt op elk moment, dus ook nadat u hebt toegestemd in deelname aan het onderzoek.

➢ **Waarom dit onderzoek?**

De laatste jaren wordt er vaker en meer onderzoek gedaan naar de structuur binnen een team en hoe dit effect heeft op het presteren (de effectiviteit) van een team.

Een team is een groep mensen die van elkaar afhankelijk zijn voor het uitvoeren van een taak, en een gezamenlijke verantwoordelijkheid hebben voor de uitkomsten (bijvoorbeeld het winnen van een wedstrijd) van dat team. Uit eerder onderzoek bleek dat bepaalde dynamieken binnen een team een positieve invloed kunnen hebben op het winnen van wedstrijden.

De Nederlandse vereniging van sportpsychologen gebruikt de Teamwork Checklist om deze processen binnen een team in kaart te brengen en te evalueren. De Teamwork Checklist meet processen binnen het team op zes verschillende dimensies op basis van bestaand onderzoek. Deze vragenlijst wordt al een aantal jaar gebruikt, maar er is nog niet gekeken naar de betrouwbaarheid van deze vragenlijst. Het doel van dit onderzoek is om naar de test-hertest betrouwbaarheid van de Teamwork Checklist te analyseren en daarvoor hebben we teams nodig die deze vragenlijst twee keer in willen vullen.

➢ **Wat vragen we van u tijdens het onderzoek?**

Voordat u deelneemt aan dit onderzoek, wordt eerst uw deelname gevraagd. Leest u deze goed door en bepaald of u deel wil nemen aan dit onderzoek. Als u goedkeuring geeft om deel te nemen aan dit onderzoek, wordt u doorgestuurd naar de teamwork checklist. Deze vragenlijst vraagt eerst wat algemene informatie van u. Daarna krijgt u een vragenlijst over uw team en de coaches. De vragen gaan over hoe het team met elkaar omgaat, zowel binnen als buiten de wedstrijden. Alle vragen worden beantwoord op een schaal van 1 tot 5.

U wordt gevraagd om de vragenlijst twee keer in te vullen met ongeveer vier tot zes weken tussen de twee invul sessies. De ingevulde vragenlijsten zullen worden geanonimiseerd en alleen het team waar u in speelt en uw leeftijd zal worden gevraagd op gebied van persoonlijke informatie. Met alle informatie die u invult zal vertrouwelijk omgegaan worden.

➢ **Welke gevolgen kan deelname hebben?**

Als u en uw team de vragenlijst invult, kunnen de resultaten gedeeld worden met uw coach mocht hiervoor interesse voor zijn en krijgt u een beter inzicht in de team dynamieken binnen uw team. Als de resultaten worden besproken met de coaches zullen er geen persoonlijke vragenlijsten gedeeld worden, enkel het gemiddelde van een team. Deze resultaten kunnen ook niet naar u als persoon worden herleid. Dit kan als resultaat hebben dat u kan werken aan mogelijke zwakke punten binnen het team en na het verbeteren van deze punten kan dat leiden tot het winnen van meer wedstrijden.

➢ **Hoe gaan we met uw gegevens om?**

Uw gegevens worden gebruikt voor het afstuderen van S. de Zeeuw. De data van uw vragenlijst wordt geanonimiseerd en gebruikt om de betrouwbaarheid van de teamwork checklist te onderzoeken. Uw naam wordt gekoppeld aan een deelnemersnummer, en deze koppeling wordt in een versleutelde database van de Rijksuniversiteit Groningen opgeslagen. Als de vragenlijst voor de tweede keer is ingevuld en uw data is gekoppeld aan uw deelnemersnummer, wordt de koppeling verwijderd en zijn de vragenlijsten alleen nog terug te herkennen aan het deelnemersnummer. Tot deze handeling is voltooid, heeft u recht om verwijdering van uw persoonsgegevens aan te vragen. Het is wettelijk verplicht om de geanonimiseerde data 10 jaar bij de Rijksuniversiteit Groningen te bewaren.

Na afronding van de these krijgt u de mogelijkheid om de publieke versie van deze these the lezen.

➢ **Wat moet u nog meer weten?**

U kunt altijd vragen stellen over het onderzoek: nu, tijdens het onderzoek, en na afloop. Dit kan door een van de betrokken onderzoekers te e-mailen (s.de.zeeuw@student.rug.nl) of te bellen ([redacted]).

Heeft u vragen/zorgen over uw rechten als onderzoeksdeelnemer of de uitvoering van het onderzoek? U kunt hierover ook contact opnemen met de Ethische Commissie Gedrags- en Maatschappijwetenschappen van de Rijksuniversiteit Groningen: ec-bss@rug.nl.

Heeft u vragen of zorgen over hoe er met uw persoonsgegevens wordt omgegaan? U kunt hierover ook contact opnemen met de Functionaris Gegevensbescherming van de Rijksuniversiteit Groningen: privacy@rug.nl.

Als onderzoeksdeelnemer heeft u recht op een kopie van deze onderzoeksinformatie.

Appendix 13: Informed consent

“Onderzoek naar de Teamwork Checklist.”

- Ik heb de informatiebrief gelezen. Ook kon ik vragen stellen. Mijn vragen zijn voldoende beantwoord. Ik had genoeg tijd om te beslissen of ik meedoe.
- Ik weet dat meedoen vrijwillig en anoniem is. Ook weet ik dat ik op ieder moment kan beslissen om toch niet mee te doen of te stoppen met de vragenlijst. Daarvoor hoef ik geen reden te geven.
- Ik weet dat ik vragen over mag slaan als ik deze niet wil beantwoorden. Ik hoef daarvoor geen reden te geven.
- Ik geef toestemming voor het verzamelen en gebruiken van mijn gegevens op de manier en voor de doelen die in de informatiebrief staan.
- Ik weet dat sommige personen mijn gegevens in kunnen zien. Die personen staan genoemd in deze informatiebrief.
- Ik geef toestemming om mijn reactie nog 15 jaar bij de Rijksuniversiteit Groningen te bewaren.