Scandinavian Journal of
Sport and Exercise Psychology

difo Dansk Idrætspsykologisk Forum

RESEARCH ARTICLE OPEN ACCESS

Psychometric Evaluation of the Norwegian Translation of the Group Conflict Questionnaire

Tommy Haugen[1], Andreas Stenling[1,2], Derek M. Peters[3], Luc Martin[4], and Rune Høigaard[1]
[1]Department of Sport Science and Physical Education, University of Agder, Kristiansand, Norway, [2]Department of Psychology, Umeå University, Umeå, Sweden, [3]Independent researcher, [4]School of Kinesiology and Health Studies, Queen's University, Kingston, Canada.

Corresponding author: Tommy Haugen
E-mail: tommy.haugen@uia.no
Department of Sport Science and Physical Education, University Agder
Post Box 422, 4604 Kristiansand S, Norway

Abstract

Perspectives on group dynamics usually acknowledge conflict in teams as unavoidable, necessitating that it be acknowledged and dealt with to develop effective and functional teams. A psychometrically sound measure is warranted to advance our understanding of this phenomenon. Based on the importance of investigating and understanding conflict in sport teams, the purpose of this study was to translate the Group Conflict Questionnaire into Norwegian using recommended scale adaptation and translation procedures and to examine the psychometric properties of the scale in Norwegian adult and youth interdependent sport team athletes. 621 recreational team-sport athletes between ages 13 and 35 were randomly split in half for exploratory factor-structure testing and a confirmatory cross-validation of the most appropriate model. The factor structure of the scale was tested using both Confirmatory Factor Analysis and Exploratory Structural Equation Modeling. Bifactor solutions were also tested. The results indicated support for a two-factor model with subsequent acceptable model-fit for the scale. The results do, however, suggest a somewhat suboptimal scale in need of additional revisions and validation, because of concerning cross-loadings and potential issues with discriminant validity. The study investigates previously untested models of the GCQ, adding valuable information about the underlying structure of the group-conflict phenomenon.

Keywords: *Group dynamics, Measurement, Norway, Sport psychology, Team sport, Bifactor model, Validation*

Highlights:

- Norwegian GCQ shows acceptable two-factor structure in athletes
- ESEM provided better model fit than traditional CFA approaches
- Evidence of strong global conflict factor via bifactor modeling
- Cross-loadings raise concerns about discriminant validity
- Reduced scale (12 items) improves model fit and clarity

In many sports, athletes must work interdependently with teammates and rely on one another to enable effective team functioning. In such environments, scholars have highlighted the inevitability that conflict will occur between members, which has implications for both individual experiences and team outcomes (DeChurch et al., 2013; Mellalieu et al., 2013; Wachsmuth et al., 2017). Indeed, several models of group development identify conflict equally unavoidable and integral, necessitating that it be acknowledged and dealt with to develop effective and functional teams (Bion, 2013; Tuckman, 1965; Miles, 1981; Wheelan, 2005). Consequently, conflict represents an omnipresent process within group dynamics that warrants continued research attention.

Barki and Hartwick (2004) defined conflict as "a dynamic process that occurs between interdependent parties as they experience negative emotional reactions to perceived disagreements and interference with the attainment of their goals" (p. 234). Gleaned from this definition is the positioning that conflict is predominantly a negative process. Indeed, Partridge and Knapp (2016) suggested that conflict could manifest itself in a variety of maladaptive or aggressive ways, resulting in an adverse impact on individual psychosocial and team levels (e.g., wellbeing, team cohesion, performance). However, researchers have also identified potential positive outcomes emanating from the resolution of conflict (Wachsmuth, et al., 2017). Therefore, the consequences of conflict are expected to vary greatly, involving problematic experiences

Corresponding author: Haugen: tommy.haugen@uia.no

of worry, confusion, frustration, disappointment, withdrawal and defensive behaviour, and more adaptive and helpful outcomes such as greater focus and task clarity, increased confidence and motivation, and proactive problem solving.

Given the inevitability of conflict in group settings, and that both maladaptive and adaptive outcomes can be expected, the negative intrapersonal consequences are well documented. Conflicts can contribute to reduced satisfaction, increased athlete isolation, stress and, competitive anxiety (Fletcher et al., 2012; Hanton et al., 2005; Olusoga et al., 2009; Paradis et al., 2014a, 2014b; Sullivan & Gee, 2007; Tamminen et al., 2013). Further, at the team level, a range of processes and emergent states can be negatively impacted, such as relationship quality, communication, coordination, collective efficacy, and cohesion (Leo et al., 2015; Paradis et al., 2014a; Sullivan & Feltz, 2001; Tamminen et al., 2013). However, if conflict is resolved effectively, athletes could more clearly understand one another, and a common solution could improve perceptions of unity and cohesion within the team (Sullivan & Feltz, 2001). Nevertheless, studies have also shown that there is a negative relation between the presence of conflict and both individual and team performance (Mellalieu et al., 2013; Partridge & Knapp, 2016).

It seems clear that it is not the presence or absence of conflict that is of importance, but rather, the way in which the conflict is engaged with and managed. Several studies have explored how different conflict management strategies can moderate the negative effects of conflict on team outcomes. These strategies are crucial in mitigating the adverse impacts of both social and task conflicts.

One study (Liu et al., 2009) found that social conflict reduces team cohesiveness and negatively affects team performance. However, using a compromising approach can help reduce these negative effects, while avoiding conflict tends to undermine team outcomes (Liu et al., 2009). Another study indicates that avoiding and integrating conflict management styles can buffer the link between social conflict and collective emotional exhaustion, whereas compromising can increase this negative link (Benitez et al., 2018).

Moreover, perceived team leader's emotional intelligence can mitigate the negative effect of social conflict on members' work engagement (Yuan et al., 2019). Conversely, members' expressive suppression strategies can intensify the negative effects of social conflict (Yuan et al., 2019). Additionally, cognitive reappraisal, an emotion regulation strategy, can help teams overcome early-stage conflict by addressing underlying appraisals of threat, ultimately preventing a decline in team performance (Thiel et al., 2019). The importance of the conflict issue itself can moderate the association between task and social conflict. When teams engage in conflicts over important task issues, the negative association between task and social conflict is diminished, partly due to a decrease in negative emotions during these conflicts (Rispens, 2012).

Even though conflict has theoretically been elaborated and empirically investigated in a variety of settings (e.g., Barki & Hartwick, 2004; Deutsch, 1990; Jehn, 1995), there is a dearth of empirical research in the context of sport (Martin et al., 2014; Paradis et al., 2014a; Wachsmuth et al., 2017). Paradis et al. (2014b) attributed this lack of investigation to the absence of a validated measurement tool, and subsequently developed the Group Conflict Questionnaire (GCQ). The GCQ is based on Barki and Hartwick's (2004) conceptualization of intra-group conflict, comprising dimensions of task (i.e., differences in viewpoints, ideas, and opinions about the task) and social conflict (i.e., issues resulting from interpersonal incompatibilities including status and commitment conflicts). Further, the authors positioned conflict as extending beyond experiences of 'disagreements' when the following three components are present: cognitive disagreement, behavioral interference with goal attainment, and affective negative emotion. Cognitive disagreement refers to differing views or ideas among individuals or groups, which can lead to misunderstandings and intellectual conflict. For example, two team-members may interpret information differently or disagree on how to address a problem. Behavioral interference with goal attainment involves actions that obstruct the achievement of goals. For instance, if one team-member does not complete their tasks, it can obstruct the entire team's progress. Affective negative emotions encompass the emotional responses during conflicts, such as frustration, anger, or stress. These feelings can arise from cognitive disagreements or behavioral issues and may escalate the conflict. For example, an intense argument might leave athletes feeling discouraged.

To ensure that the GCQ was suitable for the context of sport, Paradis et al. (2014a) first explored the way in which athletes understood and experienced intra-group conflict. Based on these findings, they generated items and assessed content validity and factorial validity through confirmatory factor analyses of the GCQ (see Paradis et al., 2014b). The two-dimensional (task and social conflict) 14-item version of the GCQ demonstrated moderate-to-acceptable model fit, and these analyses provided initial support for factorial validity, convergent validity, known-group difference validity, and discriminant validity. The authors concluded that the GCQ was promising but required further psychometric testing (Paradis et al., 2014b). Importantly, a key component of psychometric testing is to ensure that the foundation of a questionnaire can be translated for use across different contexts and cultures. To the best of our knowledge, the only translated version of the GCQ that are psychometrically tested to date is a Portuguese version on Brazilian athletes (Nascimento Junior et al., 2020).

Because of the potential negative consequences of intra-team conflict in sport, and in recognizing the dearth of empirical research in this area internationally - inclusive of the Norwegian context - a psychometrically sound measure is warranted to advance our understanding of this phenomenon. In the process of developing a language specific version of an instrument, the quality of the translation process plays a significant role (Clark & Watson, 2019). Indeed, a comprehensive and precise translation process ensures that the results obtained in one culture are not due to linguistics or contextual errors during translation, but rather, are due to real differences or similarities in the phenomena being measured.

Based on the importance of investigating and understanding conflict in sport teams, the purpose of this study was to translate the GCQ items into Norwegian using recommended scale adaptation and translation procedures and to examine the psychometric properties of the adapted scale in Norwegian adult and youth interdependent sport team athletes.

Method

The study was approved by The Norwegian Social Science Data and the Faculty ethical board at the first author's University. The study is a cross-sectional survey, using a web-based software for data-collection (SurveyXact,

Rambøll Management Consulting). All participants were recruited through secondary schools, high schools, and universities in Norway with initial contact made via their teacher/coach. Participation was voluntary and required informed consent. For participants under the age of 15, parental consent was also obtained. Although the participants were recruited from educational institutions, they completed the questionnaire in relation to their participation in competitive sport teams (outside of school).

Translation process. The GCQ items were translated into Norwegian using a process of back-translation, based on Brislin's (1970) recommendation. The English version was first translated into Norwegian by one professional translator and one experienced bilingual sport researcher. Expert reviewers consisting of one Norwegian football coach (professional soccer license) and two sport researchers with previous elite sport participation experience (PhD in Sport Science, one former elite soccer player and coach and one former elite distance runner and coach) then completed the translated versions while using a think-aloud approach. These individuals were not involved in the study, and the process was used to understand the participant thought process and to improve the clarity of the items (Jääskeläinen, 2010). The Norwegian version of the scale (GCQ-N) was then back-translated to English by an independent bilingual researcher (PhD) who was unfamiliar with the research topic, and a native English-speaking researcher (PhD) evaluated the back-translated version against the original English version. The content and meaning of each item were evaluated as entirely similar and therefore, no additional adjustments were required. The translated questionnaires were completed by six appropriately aged Norwegian interdependent sport team athletes (aged 14-28 years) who confirmed the clarity of both the instructions and the individual items. No further revisions were made after this pilot testing (see Table 1).

Participants and procedures. A convenience sample consisting of 621 athletes (age range 13-35; M_{age} = 17.0 years, *SD* = 2.84; 30% female) from different competitive teams playing football, handball, ice-hockey, and volleyball completed the questionnaire. The athletes had participated in their sport for an average of 10.11 years (*SD* = 3.47) and had been on their current teams for 5.62 years (*SD* = 4.25). The sample was randomly split in half for initial factor-structure testing (Sample 1) and cross-validation of the most appropriate model (Sample 2). The random split yielded the following samples: Sample 1 comprised 312 individuals, M_{age} = 16.8 years (*SD* = 2.51), 30% females; Sample 2 comprised 309 participants, M_{age} = 17.1 years (*SD* = 3.12), 30% females.

Measures. The Norwegian translated and adapted version of the GCQ was used (Paradis, et al., 2014b). The instrument has 14 items grouped into two factors: task conflict and social conflict (see Table 1). Agreement with items is rated on a 9-point Likert scale, with endpoint-labels of 1 (*Strongly disagree*) and 9 (*Strongly agree*). Thus, higher scores reflected stronger perceptions of intra-group conflict.

Table 1: Item list for the Group Conflice Questionnaire and the Norwegian version (GCQ-N)

Task conflict items	**Oppgave konflikt utsagn**
1. The team's ability to be successful is jeopardised because of heated disagreements during competition	1. Lagets mulighet for suksess er i fare på grunn av opphetede uenigheter under konkurranse
2. Strong disagreements during practice between members of our team disrupt our progress towards achieving team goals	2. Sterk uenighet mellom utøvere under trening ødelegger for fremgang og forbedring
3. It is nearly impossible to function effectively because of the intensity of the disagreements between members of our team during practices	3. Det er nesten umulig å fungere effektivt på trening på grunn av intense uenigheter mellom utøverene
4. The extreme animosity associated with the disagreements among members of our team affects our performance	4. Den fiendtlige stemningen på grunn av uenigheter mellom utøverne påvirker lagets prestasjon
5. The anger associated with the disagreements among members of our team affects our performance	5. Sinne knyttet til uenigheter mellom utøverne påvirker vår prestasjon
6. Members of our team have intense disagreements to the point of dysfunction	6. Spillere er så dypt uenige at laget ikke fungerer
7. There is tension among members of our team over disagreements about performance expectations	7. Det er spenning mellom utøverne på laget grunnet uenighet om forventninger til hva vi kan/skal prestere
Social conflict items	**Sosial konflikt utsagn**
8. Personal friction among members of our team leads to angry confrontations at social gatherings	8. Personlige konflikter mellom utøverne på laget ender med sinte konfrontasjoner i sosiale sammenhenger
9. The heated disagreements among members of our team in social situations become personal	9. Opphetede uenigheter mellom utøverne har blitt personlige

10. Members of our team stop speaking to each other over personal disagreements in social situations	10. Det er utøvere på laget som har sluttet å snakke til hverandre i sosiale situasjoner, på grunn av personlig uenighet
11. Emotions run high in social situations about personal differences brought to light	11. Sterke følelser settes i sving når personlige forskjeller blir diskutert
12. The negativity from personal disagreements makes it difficult for members of our team to be friends	12. Negativiteten fra personlig uenighet gjør det vanskelig for utøvere på laget å være venner
13. Members of our team have negative emotional confrontations that hinder the enjoyment of social events	13. Utøvere på laget har negative konfrontasjoner som ødelegger stemningen i sosiale sammenkomster
14. Disagreements at social gatherings escalate quickly that damages and tears our team apart	14. Uenigheter på sosiale sammenkomster eskalerer raskt, og ødelegger/splitter laget vårt

Statistical Analyses

To estimate the measurement models, Mplus version 8.4 was employed (Muthén & Muthén, 1998-2017). We used the robust full-information maximum likelihood estimator (MLR) due to its robustness to non-normality (Satorra & Bentler, 1994) and ability to account for item-level missing data (Enders, 2010). An exploratory factor analysis (EFA) with Geomin rotation was used to initially evaluate the translated scale's factor structure in Sample 1. The factor structure of the GCQ-N was then confirmed in Sample 2 using both Independent Clusters Model Confirmatory Factor Analysis (ICM-CFA) and Exploratory Structural Equation Modeling (ESEM). The latent variables of the GCQ-N were expected to be positively correlated given the results of earlier studies with the instrument (e.g., Paradis et al., 2014b). Factor correlations in ICM-CFA models are often inflated due to the restrictive nature of the model specification (i.e., no cross-loadings; Marsh et al., 2014). ESEM allows for the inclusion of cross-loadings between items and non-target factors, which has been shown to result in less biased and more accurate latent factor correlations (Marsh et al., 2014). Cross-loadings in measurement instruments, such as the GCQ, may be justified by substantive theory, item content, or may just be another source of measurement error. Items are not perfect representations of the constructs and often have weak residual relations with other constructs (Asparouhov & Muthén, 2009). It is often reasonable to assume non-zero cross-loadings in measures of psychological constructs because most items have several determinants (Marsh et al., 2014).

Additionally, simulations and empirical studies demonstrate that forcing minor cross-loadings to zero or ignoring them altogether may result in inflated factor correlations that reduce discriminant validity and produce inaccurate estimates in SEM when combined with other variables (Marsh et al., 2014). We used target rotation in the ESEM (Asparouhov & Muthén, 2009), whereby all cross-loadings were set to be close to zero but not exactly zero, and the target loadings were freely estimated. The latent factors in the ICM-CFA and ESEM models were scaled using the fixed-factor approach.

Target factor loadings of 0.500 or above were regarded as suitable, whereas values below 0.300 denoted an insufficient indicator (Morin et al., 2020). Target loadings between 0.300 and 0.500 should be carefully examined in further research. Cross-loadings below 0.300 were seen as negligible in the current study, whereas cross-loadings over 0.300 were regarded as noteworthy and warranted closer examination.

Multiple goodness-of-fit indices were used to evaluate model fit. The Tucker Lewis index (TLI), comparative fit index (CFI), root mean square error of approximation (RMSEA), and standardized root mean square residual (SRMR) were evaluated, more explicitly based on the following criteria: TLI and CFI >.90; RMSEA < .08, and SRMR < .08 (Marsh, 2007). A difference of less than 0.010 on the CFI and 0.015 on the RMSEA between two models can be regarded as proof that they offer a comparable fit to the data for nested model comparisons (i.e., the ICM-CFA vs. ESEM; Chen, 2007). Using standardized parameter estimates from the confirmatory models, omega reliability (ω) was calculated in accordance with McDonald's (1970) formula, $\omega = (\Sigma|\lambda i|)^2/([\Sigma|\lambda i|^2]+\Sigma\delta i)$ where λi represents the factor loadings and δi represents the error variances (Geldhof et al., 2014). Based on the magnitude of the correlations between the latent variables, researchers have proposed numerous criteria for establishing discriminant validity (Cheung & Wang, 2017; Kline, 2015). In the current study, correlations higher than 0.70 were viewed as an indication of problems with discriminant validity (Cheung & Wang, 2017). Improved model fit, attenuated latent factor correlations, strong target factor loadings, and relatively weak cross-loadings in the ESEM compared to the ICM-CFA were used as evidence that ESEM fits the data better than the ICM-CFA. Note that we also report results from the chi-square test of exact fit, but we do not rely on it for model fit evaluation due to its oversensitivity to sample size and minor and substantively unimportant misspecifications (Marsh et al., 2014).

Based on an assumption that there may be a common underlying construct (i.e., group conflict) but also specific conflict dimensions (i.e., task conflict and social conflict) we also examined a bifactor solution of the GCQ-N. A bifactor model is designed to test the existence of a general factor along with specific factors that are unique to their own sub-dimensions (Morin et al., 2020). In a bifactor model, variance in the items is explained by three components: a general factor, a specific factor, and residual variance. The general factor represents the common variance shared by all test items, whereas the specific factors capture the unique variance specific to each subset of items. The bifactor model allows for the estimation of the effects of the general factor on the observed variables, as well as the effects of the specific factors on their respective observed variables (Morin et al., 2020). Both models (two-factor solution and bifactor solution) were tested within the framework of ICM-CFA and ESEM. The data and syntax can be obtained through the Open Science Framework (https://doi.org/10.17605/OSF.IO/NJXYR).

Results

Table 2 presents descriptive statistics for the 14 items in the GCQ-N. Skewness and kurtosis values were within the range of -2 and 2 and the mean values of the items were at the lower end of the response scale (< 3.60).

Table 2: Descriptive Statistics of the Items from the GCQ-N

	Sample 1					
	M	*SD*	Skewness	Kurtosis	Range	*n*
TCT1	3.623	5.158	0.540	-0.719	1-9	310
TCT2	3.564	3.564	0.676	-0.564	1-9	307
TCT3	2.705	4.156	1.238	0.762	1-9	308
TCT4	3.033	4.858	0.901	-0.253	1-9	305
TCT5	3.214	5.148	0.770	-0.549	1-9	304
TCT6	2.379	4.262	1.539	1.387	1-9	306
TCT7	3.058	4.406	0.900	-0.060	1-9	308
TCS8	2.587	4.065	1.211	0.538	1-9	305
TCS9	2.771	4.510	1.091	0.197	1-9	306
TCS10	2.696	5.150	1.257	0.524	1-9	309
TCS11	2.829	4.813	1.166	0.346	1-9	304
TCS12	2.788	5.014	1.110	0.091	1-9	307
TCS13	2.633	4.383	1.285	0.843	1-9	305
TCS14	2.516	5.087	1.509	1.153	1-9	308
	Sample 2					
	M	*SD*	Skewness	Kurtosis	Range	*n*
TCT1	3.536	5.125	0.579	-0.618	1-9	306
TCT2	3.430	5.489	0.797	-0.379	1-9	307
TCT3	2.711	4.069	1.318	1.120	1-9	308
TCT4	3.000	5.320	1.118	0.172	1-9	306
TCT5	3.134	5.408	0.901	-0.317	1-9	305
TCT6	2.299	3.769	1.579	1.532	1-9	304
TCT7	2.944	4.265	0.847	-0.397	1-9	302
TCS8	2.524	4.184	1.264	0.558	1-9	307
TCS9	2.535	3.861	1.204	0.333	1-9	299
TCS10	2.306	4.206	1.562	1.310	1-9	304
TCS11	2.558	4.519	1.361	0.840	1-9	301
TCS12	2.455	4.162	1.456	1.202	1-9	301
TCS13	2.447	4.234	1.493	1.359	1-9	304
TCS14	2.249	3.853	1.736	2.254	1-9	305

Note. TCT= team conflict task, TCS = team conflict social

Table 3 presents the goodness of fit statistics of the EFA on Sample 1. The three-factor solution yielded a marginally better model fit compared to the two-factor solution. However, the third factor in the three-factor solution had only one statistically significant factor loading. Thus, the three-factor solution was not interpreted as conceptually meaningful, and the two-factor solution was deemed most appropriate. The factor-correlations between the two factors in the EFA was 0.79.

Table 3: Goodness of Fit Statistics and Information Criteria for the Models Estimated on the GCQ-N Scale

Model	χ^2	df	CFI	TLI	RMSEA	RMSEA 90% CI	SRMR	AIC	BIC	ABIC
EFA (sample I)										
One factor solution	374.359*	77	0.863	0.838	0.111	[0.100, 0.123]	0.053	14973	15130	14997
Two factor solution	181.462*	64	0.946	0.923	0.077	[0.064, 0.090]	0.027	14587	14793	14618
Three factor solution	143.158*	52	0.958	0.926	0.075	[0.060, 0.090]	0.024	14519	14769	14557
ICM-CFA (sample II)										
Two factor solution	248.334*	76	0.924	0.909	0.086	[0.074, 0.098]	0.052	14268	14429	14292
Bifactor solution	153.376*	63	0.960	0.942	0.068	[0.055, 0.082]	0.026	14120	14329	14151
ESEM (sample II)										
Two factor ESEM	166.822*	64	0.955	0.935	0.072	[0.059, 0.086]	0.024	14162	14367	14193
Bifactor ESEM	101.381*	52	0.978	0.962	0.055	[0.039, 0.071]	0.017	14088	14338	14126

Note. df = Degrees of freedom; CFI = comparative fit index; TLI = Tucker-Lewis index; RMSEA = root mean square error of approximation; CI = confidence interval; SRMR = standardized root mean square residual; AIC = Akaike information criterion; CAIC = Constant AIC; BIC = Bayesian information criterion; ABIC = Sample size adjusted BIC. * All χ^2 values are significant ($p < .001$).

As presented in Table 4, when examining the two-factor solution, results showed that all task items loaded statistically significant on factor 1 (factor loadings ranging from 0.557 to 0.940) and the social items loaded statistically significant on factor 2 (factor loadings ranging from 0.584 to 0.952). Two notable cross-loadings were identified in the EFA (TCT6 on factor 2, 0.319; TCS8 on factor 1, 0.328). However, both items loaded > 0.500 on their intended factor and were thus retained in the ICM-CFA and ESEM on Sample 2. As can be seen in Table 3, the two-factor solution of the GCQ-N was confirmed with acceptable model fit in the ICM-CFA. The two-factor ESEM did, however, yield better fit compared to the ICM-CFA model. Omega reliability estimates based on the confirmatory analysis were 0.934 for the task conflict dimension and 0.960 for the social conflict dimension; and based on the ESEM-model 0.916 for task conflict dimension and 0.943 for the social conflict dimension.

Despite better fit for the ESEM-model, an inspection of the factor loadings presented in Table 5 showed that one item (TCT4) had a standardized factor loading > 1.000. Moreover, one item had a higher factor loading on the non-intended factor compared to the intended factor (TCT6: 0.417 on TCT and 0.477 on TCS). One additional cross-loading above the cut-off (> 0.300) appeared for TCT7 (0.311). The latent factor correlations (TCS with TCT) were 0.83 for the ICM-CFA and 0.79 for the ESEM.

Table 4: Factor loadings for EFA (Sample 1)

Items	One Factor	Two Factors		Three Factors		
	F1	F1	F2	F1	F2	F3
TCT1	0.645*	**0.557***	0.131	**0.545***	0.150	-0.006
TCT2	0.723*	**0.654***	0.122	**0.640***	0.136	0.006
TCT3	0.829*	**0.818***	0.079	**0.747***	-0.013	0.176
TCT4	0.849*	**0.940***	-0.008	**0.953***	0.003	-0.019
TCT5	0.802*	**0.906***	-0.020	**0.902***	-0.017	0.008
TCT6	0.840*	**0.572***	0.319	0.457	0.119	0.336
TCT7	0.821*	**0.609***	0.266	0.549	0.191	0.145
TCS8	0.875*	0.328	**0.584***	-0.023	-0.001	**1.013***
TCS9	0.837*	0.257	**0.617***	0.084	0.318	0.502
TCS10	0.802*	-0.122	**0.952***	-0.131	**0.947***	0.017
TCS11	0.865*	0.006	**0.898***	-0.024	**0.860***	0.075
TCS12	0.899*	0.104	**0.839***	0.107	**0.895***	-0.057
TCS13	0.860*	-0.022	**0.918***	-0.016	**0.925***	-0.013
TCS14	0.858*	0.107	**0.785***	0.053	**0.691***	0.158

Note. N = 314. TCT = team conflict task, TCS = team conflict social * = $p < .01$

The bifactor models provided good fit to the data (see Table 3), with the ESEM-model superior to the ICM-CFA. The factor loading pattern from the bifactor ESEM shows that all items loaded relatively strongly onto the global conflict factor (standardized factor loadings > .656). However, the specific factors, particularly task conflict, were relatively poorly defined (Table 5).

Table 5: Standardized Factor Loadings and Item Uniquenesses from ICM-CFA and ESEM Models of the GCQ-N Scale (Sample 2)

	ICM-CFA			ESEM			B-ICM-CFA				B-ESEM			
	TCT	TCS	δ	TCT	TCS	δ	TCT	TCS	TC	δ	TCT	TCS	TC	δ
TCT1	**0.716***		0.487*	**0.721***	-0.011	0.493*	**0.387***		**0.596***	0.495*	**0.337***	0.001	**0.669***	0.439*
TCT2	**0.805***		0.351*	**0.857***	-0.072	0.358*	**0.469***		**0.647***	0.362*	**0.668***	-0.026	**0.737***	0.009
TCT3	**0.873***		0.238*	**0.848***	0.023	0.250*	**0.451***		**0.737***	0.253*	**0.195***	-0.017	**0.842***	0.252*
TCT4	**0.913***		0.166*	**1.032***	-0.126*	0.123*	**0.577***		**0.742***	0.117*	**0.032**	-0.179	**0.940***	0.083*
TCT5	**0.894***		0.200*	**0.969***	-0.078*	0.174*	**0.536***		**0.737***	0.169*	**0.040**	-0.127	**0.906***	0.161*
TCT6	**0.814***		0.337*	**0.417***	0.477*	0.285*	**0.143***		**0.849***	0.259*	**0.001**	0.236*	**0.812***	0.285*
TCT7	**0.758***		0.425*	**0.490***	0.311*	0.424*	**0.182***		**0.753***	0.400*	**0.153**	0.173*	**0.723***	0.424*
TCS8		**0.857***	0.266*	0.287*	**0.623***	0.248*		**-0.043**	**0.912***	0.166*	0.024	**0.328***	**0.802***	0.248*
TCS9		**0.919***	0.155*	0.091*	**0.846***	0.154*		**0.140**	**0.916***	0.142*	-0.028	**0.456***	**0.799***	0.152*
TCS10		**0.825***	0.320*	-0.138	**0.948***	0.287*		**0.249***	**0.787***	0.318*	-0.042	**0.528***	**0.657***	0.287*
TCS11		**0.826***	0.319*	0.028	**0.805***	0.315*		**0.255***	**0.785***	0.318*	-0.037	**0.433***	**0.705***	0.314*
TCS12		**0.919***	0.155*	0.016	**0.907***	0.155*		**0.405***	**0.845***	0.122*	0.033	**0.508***	**0.766***	0.154*
TCS13		**0.920***	0.153*	0.047	**0.881***	0.156*		**0.386***	**0.851***	0.127*	0.061	**0.495***	**0.774***	0.152*
TCS14		**0.901***	0.188*	0.090	**0.826***	0.194*		**0.285***	**0.851***	0.195*	-0.021	**0.440***	**0.783***	0.193*

Note. TCT = team conflict task, TCS = team conflict social, TC = team conflict, δ = item uniquenesses (residual variances).
*$p < .05$. Latent factor-correlations: TCS with TCT: CFA = 0.825, ESEM = 0.787

Post-hoc modifications of the scale

Because of the challenges with all the tested models (e.g., poor model fit, substantial cross-loadings, standardized factor loading > 1.000) we decided to explore the model-fit of a reduced version of the scale, similar to the version found in Nascimento Junior et al. (2020), with post-hoc analyses on the same sub-sample (sample 2). We opted to remove the two items with concerning cross-loadings (TCT6 and TCT7). Both the two-factor ICM-CFA and ESEM were tested. Results from the additional analyses are shown in appendix 1 (Tables A1 and A2). The result revealed acceptable model-fit for the ICM-CFA (S-B $\chi2$ = 139.465 [*df* = 53, *N* = 309], $p < .0.001$; CFI = .953; TLI = .941; RMSEA = .073 [.058 – .087], and SRMR = .033), with standardized factor loadings between 0.714 and 0.936. Latent factor-correlation was 0.78 ($p < 0.001$). The ESEM-model also yielded acceptable model-fit (S-B $\chi2$ = 111.140 [*df* = 43, *N* = 309], $p < .0.001$; CFI = .963; TLI = .943; RMSEA = .072 [.055 – .088], and SRMR = .023). No notable cross-loadings were identified. The latent factor correlation for the ESEM was 0.77 ($p < 0.001$).

Discussion

The purpose of the present study was to translate the GCQ into Norwegian and examine the psychometric properties of the scale on Norwegian team sport athletes. The study followed a thorough translation process in line with recommendations in the literature. The initial EFA in Sample 1 indicated that a two-factor solution was the most appropriate, and the items loaded acceptably onto their intended factors. Contrary to the Portuguese version of the scale (Nascimento Junior et al., 2020) no items were removed after the initial analysis. Testing both ICM-CFA and ESEM in Sample 2 indicated acceptable fit for both models, however the ESEM was deemed superior to the ICM-CFA because of the better model-fit.

The latent-factor correlation between task and social conflict in the present study mirrored the findings from the phase 2[1] in Paradis et al. (2014b) and could be considered relatively strong (ICM-CFA = 0.83; ESEM = 0.79), indicating possible issues with discriminant validity. In contrast, Nascimento Junior et al. (2020) found a latent factor correlation of 0.63 in their Portuguese version. It should be noted, however, that the Portuguese version omitted the two items that in the present study had the highest cross-loading (TCT6 and TCT7). In addition, the ICM-CFA in Nascimento Junior et al. (2020) included cross-loadings between two items on the task dimension and two items on the social dimension, which could have reduced the latent factor correlation. Thus, there could be some concerns about the discriminant validity of the GCQ, possibly across cultures/languages, and these may be more pronounced in the Norwegian version compared to other language versions of the GCQ. The composite reliability of the two sub-dimensions was deemed acceptable, similarly to the findings from Paradis et al. (2014b) and Nascimento Junior et al. (2020).

Two items intended to target the sub-dimension of task-conflict (TCT6, TCT7) had notable cross-loadings in the ESEM. Indeed, the TCT6 had higher factor loading on the social-conflict sub-dimension. These somewhat problematic cross-loadings closely resemble those from Nascimento Junior et al. (2020), who decided to exclude these two items during their exploratory phase. The initial EFA in the present study did not identify problems with the two items (see Table 2), and the CFA yielded acceptable model fit and strong factor loadings, which is similar to the original work by Paradis et al. (2014b). However, the cross-

[1] Interestingly, the phase 3 of Paradis et al. (2014) found a bivariate latent-factor correlation of 0.65, which is considerably lower than what is presented from phase 2.

loadings identified in the ESEM highlights potential problems with these items that should be considered in future psychometric testing of the GCQ. Looking closely at the wording of the item TCT6 ("members of our team have intense disagreements to the point of dysfunction"), it seems reasonable to question whether these "disagreements" are indeed perceived as task-related, as intended (Paradis et al., 2014b). Moreover, one may argue that the item is 'double barrelled', because individuals must consider (1) if members have intense disagreements and (2) if it causes dysfunction. Similar ambiguity may be apparent in the item TCT7 ("there is tension among members of our team over disagreements about performance expectations"). The item may reflect a situation where disagreements about performance expectations are not just about the tasks themselves but also tension that affect the social dynamics within the team.

As previously noted, the ESEM of the two-dimensional 14-item scale yielded acceptable model fit. However, one item had a standardized factor loading larger than 1.0 (TCT4). Jöreskog (1999) has previously argued that the factor loadings should be considered as regression coefficients and not correlations, and thus, their magnitude can be larger than 1.0. Despite this, findings such as these might indicate problems in the data (e.g., multicollinearity), however, no such problems were identified in the current study.

The model-fit of the bifactor models, particularly the ESEM, could also be considered adequate. All items loaded strongly onto the global conflict factor, which suggests that the global factor explains a substantial amount of the covariance among the items. These findings suggest that the GCQ-N captures a strong common conflict factor but there seems to be indications of multidimensionality caused by clusters of items representing the subdomains of task and social conflict (Reise, 2012). Thus, if researchers are primarily interested in capturing the global phenomenon of conflict, such a bifactor model may be relevant.

Because of the concerns raised by the initial testing of the scale, we decided to perform additional (post-hoc) tests on a reduced scale. Although the ICM-CFA and the ESEM models of the reduced version of the scale (when excluding TCT6 and TCT7, as in Nascimento Junior et al., 2020) yielded acceptable model-fit indices, one cannot directly compare these models to the original models tested, because models are no longer nested when removing items. Thus, claims about "best model fit" should be considered speculative. One may, however, use the additional analyses as a guide for future studies, encouraging further research on the psychometric properties of the GCQ. Although the post-hoc modifications made in the present study (removing items and performing additional tests on the same sample) are normally discouraged in the literature (see e.g., Pan et al., 2017; Steiger, 1990), the present finding alongside the findings from Nascimento Junior et al. (2020), clearly identify issues with these two items. We thus suggest that future research should consider removing or rephrasing these items before subsequent psychometric testing.

Conclusions and future research

As with all scientific research, this study must be read through the lenses of its strengths and limitations. Overall, the present study presents a translated version in a new language (Norwegian) and contributes to further scrutinizing of the psychometric properties of the GCQ-scale. The present study investigates previously untested models of the GCQ, adding valuable information about the underlying structure of the group-conflict phenomenon.

Two substantial limitations regarding the present study should be explicitly addressed. Firstly, the present study cannot account for the possible clustered nature of the collected data. The sample(s) consists of participants belonging to different teams. Even though the collection took place at high schools and universities, there is a chance that some participants belonged to the same teams. We have no information on team level and thus cannot account for potential clustering-effects. This should be considered when interpreting the present findings, particularly because the phenomenon of interest (team conflict) is, by definition, connected to within-group dynamics. Secondly, the present study would have benefitted by a more thorough investigation of the nomological network (Cronbach & Meehl, 1955) of team conflict. The results suggest a somewhat suboptimal scale in need of additional revisions and validation. As previously noted, the post-hoc analyses performed were done on the same sample as the original confirmatory analyses. This should be considered less than optimal, and future studies on independent samples are needed for future guidance of the psychometric properties of the scale. One approach that could be adopted in future studies on the GCQ is Bayesian CFA (B-CFA; see e.g., Muthén & Asparouhov, 2012; Stenling et al., 2015) that can incorporate cross-loadings and residual covariances among all items. Testing the CGQ using prior specifications on the residual variances and covariances might provide additional insight into the factor structure of the scale. In addition, the present samples include athletes younger than the original validation from Paradis et al. (2014b), and although the Portuguese version previously has been utilized on similar aged athletes (e.g., Freire et al., 2022), future studies would do well to investigate potential age-related differences in the psychometrics of the scale.

References

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

Barki, H., & Hartwick, J. (2004). Conceptualizing the construct of interpersonal conflict. *International Journal of Conflict Management, 15*, 216–244. https://doi.org/10.1108/eb022913

Benitez, M., Medina, F., & Munduate, L. (2018). Buffering relationship conflict consequences in teams working in real organizations. *International Journal of Conflict Management, 29*, 279-297. https://doi.org/10.1108/IJCMA-11-2017-0131

Bion, W. R. (2013). *Experiences in groups: and other papers*. Routledge. https://doi.org/10.4324/9780203359075

Brislin, R. W. (1970). Back-translation for cross-cultural research. *Journal of cross-cultural psychology*, *1*(3), 185-216. https://doi.org/10.1177/135910457000100301

Chen, F. F. (2007). Sensitivity of goodness of fit indexes to lack of measurement invariance. *Structural Equation Modeling: a Multidisciplinary Journal*, *14*(3), 464-504. https://doi.org/10.1080/10705510701301834

Cheung, G. W., & Wang, C. (2017). Current approaches for assessing convergent and discriminant validity with SEM: Issues and solutions. In *Academy of Management Proceedings* (Vol. 2017, No. 1, p. 12706). Academy of Management, Briacliff Manor, NY 10510. https://doi.org/10.5465/AMBPP.2017.12706abstract

Clark, L. A., & Watson, D. (2019). Constructing validity: New developments in creating objective measuring instruments. *Psychological Assessment, 31*(12), 1412–1427. https://doi.org/10.1037/pas0000626

Cronbach, L. J., & Meehl, P. E. (1955). Construct validity in psychological tests. *Psychological Bulletin*, *52*(4), 281-302. https://doi.org/10.1037/h0040957

DeChurch, L. A., Mesmer-Magnus, J. R., & Doty, D. (2013). Moving beyond relationship and task conflict: Toward a process-state perspective. *Journal of Applied Psychology, 98*, 559–578. https://doi.org/10.1037/a0032896

Deutsch, M. (1990). Sixty years of conflict. *International Journal of Conflict Management, 1*, 237-263. https://doi.org/10.1108/eb022682

Enders, C. K. (2010). *Applied missing data analysis*. Guilford Press.

Fletcher, D., Hanton, S., Mellalieu, S. D., & Neil, R. (2012). A conceptual framework of organizational stressors in sport performers. *Scandinavian Journal of Medicine & Science in Sports, 22*(4), 545–557. https://doi.org/10.1111/j.1600-0838.2010.01242.x

Freire, G. L. M., Fiorese, L., Moraes, J. F. V. D., Codonhato, R., Oliveira, D. V. D., & Nascimento Junior, J. R. D. A. D. (2022). Do perfectionism traits predict team cohesion and group conflict among youth athletes?. *Perceptual and Motor Skills*, *129*(3), 851-868. https://doi.org/10.1177/00315125221087025

Geldhof, G. J., Preacher, K. J., & Zyphur, M. J. (2014). Reliability estimation in a multilevel confirmatory factor analysis framework. *Psychological Methods, 19*(1), 72-91. https://doi.org/10.1037/a0032138

Hanton, S., Fletcher, D., & Coughlan, G. (2005). Stress in elite sport performers: A comparative study of competitive and organizational stressors. *Journal of Sports Sciences, 23*(10), 1129–1141. https://doi.org/10.1080/02640410400023231

Jääskeläinen, R. (2010). Think-aloud protocol. In Y. Gambier & L. Van Doorslaer (Eds.), *Handbook of translation studies* (pp. 371–374). John Benjamins Publishing https://doi.org/10.1075/hts.1.thi1

Jöreskog, K. G. (1999). *How large can a standardized coefficient be?* In the help-file of the LISREL program. Scientific Software International.

Jehn, K. A. (1995). A multi-method examination of the benefits and detriments of intragroup conflict. *Administrative Science Quarterly, 40*, 256–282. https://doi.org/10.2307/2393638

Kline, R. B. (2015). *Principles and practice of structural equation modeling*. Guilford publications.

Liu, J., Fu, P., & Liu, S. (2009). Conflicts in top management teams and team/firm outcomes. *International Journal of Conflict Management, 20,* 228-250. https://doi.org/10.1108/10444060910974867

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

Marsh, H. W., Morin, A. J., Parker, P. D., & Kaur, G. (2014). Exploratory structural equation modeling: An integration of the best features of exploratory and confirmatory factor analysis. *Annual Review of Clinical Psychology, 10*(1), 85–110. https://doi.org/10.1146/annurev-clinpsy-032813-153700

Martin, L., Bruner, M., Eys, M., & Spink, K. (2014). The social environment in sport: Selected topics. *International review of sport and exercise psychology, 7*(1), 87-105. https://doi.org/10.1080/1750984X.2014.885553

McDonald, R. P. (1970). The theoretical foundations of principal factor analysis, canonical factor analysis, and alpha factor analysis. *British Journal of Mathematical and Statistical Psychology*, *23*(1), 1-21. https://doi.org/10.1111/j.2044-8317.1970.tb00432.x

Miles, M. B. (1981) *Learning to work in groups. A practical guide to for members and trainers*. Teachers College Press.

Mellalieu, S., Shearer, D. A., & Shearer, C. (2013). A preliminary survey of interpersonal conflict at major games and championships. *The Sport Psychologist, 27*, 120–129. https://doi.org/10.1123/tsp.27.2.120

Morin, A. J. S., Myers, N. D., & Lee, S. (2020). Modern factor analytic techniques: Bifactor models, exploratory structural equation modeling (ESEM), and bifactor-ESEM. In G. Tenenbaum & R. C. Eklund (Eds.), *Handbook of Sport Psychology* (pp. 1044–1073). Wiley. https://doi.org/10.1002/9781119568124.ch51

Muthén, B., & Asparouhov, T. (2012). Bayesian structural equation modeling: A more flexible representation of substantive theory. *Psychological Methods*, 17(3), 313–335. https://doi.org/10.1037/a0026802

Muthén, L. K., & Muthén, B. O. (1998–2017). *Mplus user's guide. Eighth Edition*. Muthén & Muthén.

Nascimento Junior, J. R. A. D., Codonhato, R., Freire, G. L. M., Fortes, L. D. S., Quinaud, R. T., Carvalho, H. M., ... & Fiorese, L. (2020). Group Conflict Questionnaire: Adaptation and evidences of validity for Portuguese-speaking athletes. *Perceptual and Motor Skills, 127*(3), 587-608. https://doi.org/10.1177/0031512520914979

Olusoga, P., Butt, J., Hays, K., & Maynard, I. (2009). Stress in elite sports coaching: Identifying stressors. *Journal of Applied Sport Psychology, 21*(4), 442-459. https://doi.org/10.1080/10413200903222921

Pan, J., Ip, E. H., & Dubé, L. (2017). An alternative to post hoc model modification in confirmatory factor analysis: The Bayesian lasso. *Psychological Methods, 22*(4), 687–704. https://doi.org/10.1037/met0000112

Paradis, K. F., Carron, A. V., & Martin, L. J. (2014a). Athlete perceptions of intra-group conflict in sport teams. *Sport & Exercise Psychology Review, 10*(3), 4–18. https://doi.org/10.53841/bpssepr.2014.10.3.4

Paradis, K., Carron, A., & Martin, L. (2014b). Development and validation of an inventory to assess conflict in sport teams: The Group Conflict Questionnaire. *Journal of Sports Sciences, 32*, 1966–1978. https://doi.org/10.1080/02640414.2014.900847

Partridge, J. A., & Knapp, B. A. (2016) Mean girls: Adolescent female athletes and peer conflict in sport. *Journal of Applied Sport Psychology, 28*, 113-127. https://doi.org/10.1080/10413200.2015.1076088

Reise, S. P. (2012). The rediscovery of bifactor measurement models. *Multivariate behavioral research*, *47*(5), 667-696. https://doi.org/10.1080/00273171.2012.715555

Rispens, S. (2012). The influence of conflict issue importance on the co-occurrence of task and relationship conflict in teams. *Applied Psychology, 61*, 349-367. https://doi.org/10.1111/j.1464-0597.2011.00473.x

Satorra, A., & Bentler, E. M. (1994). Corrections to test statistics and standard errors in covariance structure analysis. In A. Von Eye & C. C. Clogg (Eds.), *Latent variables analysis: Applications for developmental research* (pp. 399–419). Sage.

Steiger, J. H. (1990). Structural model evaluation and modification: An interval estimation approach. *Multivariate Behavioral Research, 25*(2), 173–180. https://doi.org/10.1207/s15327906mbr2502_4

Stenling, A., Ivarsson, A., Johnson, U., & Lindwall, M. (2015). Bayesian structural equation modeling in sport and exercise psychology. *Journal of Sport and Exercise Psychology*, *37*(4), 410-420. https://doi.org/10.1123/jsep.2014-0330

Sullivan, P. J., & Feltz, D. L. (2001). The relationship between intrateam conflict and cohesion within hockey teams. *Small Group Research, 32*, 342–355. https://doi.org/10.1177/104649640103200304

Sullivan, P. J., & Gee, C. J. (2007). The relationship between athletic satisfaction and intrateam communication. *Group Dynamics: Theory, Research, and Practice, 11*, 107–116. https://doi.org/10.1037/1089-2699.11.2.107

Tamminen, K. A., Holt, N. L., & Neely, K. C. (2013). Exploring adversity and the potential for growth among elite female athletes. *Psychology of Sport and Exercise, 14*, 28–36. https://doi.org/10.1016/j.psychsport.2012.07.002

Thiel, C., Harvey, J., Courtright, S., & Bradley, B. (2019). What doesn't kill you makes you stronger: How teams rebound from early-stage relationship conflict. *Journal of Management, 45*, 1623 - 1659. https://doi.org/10.1177/0149206317729026

Tuckman, B. W. (1965). Developmental sequence in small groups. *Psychological Bulletin, 63*, 384–399. https://doi.org/10.1037/h0022100

Leo, F. M., González-Ponce, I., Sánchez-Miguel, P. A., Ivarsson, A., & García-Calvo, T. (2015). Role ambiguity, role conflict, team conflict, cohesion and collective efficacy in sport teams: A multilevel analysis. *Psychology of Sport and Exercise, 20*, 60–66. https://doi.org/10.1016/j.psychsport.2015.04.009

Wachsmuth, S., Jowett, S., & Harwood, C. G. (2017). Conflict among athletes and their coaches: what is the theory and research so far? *International Review of Sport and Exercise Psychology, 10*, 84-107. https://doi.org/10.1080/1750984X.2016.1184698

Wheelan S. A. (2005). *Group Processes. A developmental perspective 2nd ed*. Pearson.

Yuan, L., Yu, Y., & Liu, P. (2019). Are teams experiencing relationship conflicts destined to fail. *Chinese Management Studies, 14,* 235-256. https://doi.org/10.1108/CMS-01-2019-0010