A Psychometric Evaluation of the Group Environment Questionnaire in a Sample of Elite and Regional Level Greek Volleyball Players

Submission Date: 11/02/04

Re-Submission Date: 07/06/04

Word Count: 6,752

Ntoumanis Nikos[1], Aggelonidis Yannis[2]

[1] School of Sport and Exercise Sciences, University of Birmingham, UK

2 Department of Physical Education and Sports Science, University of Athens, Greece

Please address all correspondence to:

Dr. Nikos Ntoumanis, School of Sport and Exercise Sciences, University of Birmingham, Birmingham, B15 2TT, Tel: 0121 4147981, Fax: 0121 4144121,

E-mail: N.Ntoumanis@bham.ac.uk

Abstract

The purpose of this study was to examine the psychometric properties of the Group Environment Questionnaire (GEQ; Carron, Widmeyer and Brawley, 1985) adapted to the Greek language. The sample consisted of 586 male and female volleyball players of elite and regional level status. Data were analysed from three time points of a competitive season. For each time point, seven competing first-order and second-order factor structures were subjected to confirmatory factor analyses. The results revealed that the Greek GEQ demonstrated high internal reliability coefficients, good convergent validity and, for most of the competing models, acceptable fit indices. However, very high factor correlations rendered problematic the discriminant validity of the questionnaire. Multisample analyses examining the invariance of the seven models across competitive level and gender revealed that the models were largely invariant. Further psychometric testing is needed to examine whether the Greek GEQ relates to conceptually important personal and team correlates of group cohesion.

Key Words: group cohesion, Greek volleyball, confirmatory factor analysis

Biographical Note

Dr. Nikos Ntoumanis has an MSc in Sports Sciences from Loughborough University and a PhD in Medical Sciences from the University of Exeter. He is a Lecturer in Sport and Exercise Psychology at the University of Birmingham in the UK.

Dr. Yannis Aggelonidis has a PhD in Social Sport Psychology from the University of Athens. He is an Associate Professor in Volleyball Coaching at the University of Athens in Greece.

Over the last two decades the concept of group cohesion has received considerably research attention in the sport psychology literature. Carron (1982) defined group cohesion as “a dynamic process which is reflected in the tendency for a group to stick together and remain united in the pursuit of goals and objectives” (p. 124). Research evidence has shown that group cohesion is positively related to important individual and group level variables such as performance, efficacy, and satisfaction (for a comprehensive review, see Carron and Hausenblas, 1998).

Research on group cohesion in sport has been predominantly based on Carron et al.’s (1985) conceptual model of group cohesion. According to this model, there are two major categories of group cohesion. The first one is called group integration and reflects athletes’ perceptions of how the group functions as a unit. The second category is labelled individual attraction to the group and reflects how attractive the group is to the individual. Each of these categories is divided into two further categories, task and social. Task cohesion reflects the degree to which the team-members work together to achieve specific team goals. Social cohesion reflects the degree to which the team-members like each other and have good social relationships. Thus, in total four dimensions of group cohesion have been proposed by Carron et al. (1985). These are: Group Integration-Social (GI-S; how the team functions at a social level), Group Integration-Task (GI-T; how the team functions to achieve important team goals), Individual Attraction to the Group-Social (ATG-S; the extent to which athletes are attracted to the team by its social environment), and Individual Attraction to the Group-Task (ATG-T; the extent to which athletes are attracted to the team to achieve important goals. Carron et al. (1985) proposed a hierarchical conceptual model with two second-order factors, ATG and GI, underpinned by the respective task and social ATG and GI first-order factors.

The measurement of group cohesion in sport has been predominantly based on the Group Environment Questionnaire (GEQ; Widmeyer, Brawley and Carron, 1985). This questionnaire consists of eighteen items that measure the four dimensions of group cohesion. Research evidence regarding the psychometric properties of the GEQ has offered mixed results. Widmeyer et al. (1985), using exploratory factor analysis, supported the hypothesised four-factor structure of the GEQ. Employing the more rigorous method of confirmatory factor analysis, Li and Harmer (1996) examined seven different competing factorial structures. The results showed that two were the most appropriate factorial structures: one that represented a four-factor structure, and a second one that posited the hierarchical structure proposed by Carron et al. (1985). However, it should be noted that none of the seven competing structures met the more recent criteria for adequate model fit proposed by Hu and Bentler (1999). These criteria are described in the Method of our paper. Li and Harmer (1996) also argued that the GEQ factors had acceptable convergent validity because the items loaded moderately highly on their hypothesised factors. In contrast, the discriminant validity of the GEQ was relatively poor because the four cohesion factors were highly correlated (correlations ranged from r = .71 to r = .91; average r =. 82). Although Li and Harmer (1996) showed that the factor correlations were significantly different from unity, one cannot ignore the fact that the correlations were substantially high, casting doubt on the discriminant validity of the GEQ.

Other studies have not supported the factorial structure of the GEQ. Schutz, Eom, Smool and Smith (1994) also examined a number of competing factor structures using confirmatory factor analysis. Their findings failed to support any of the competing first-order and hierarchical models. Furthermore, none of the subscales had Cronbach’s alpha internal reliability coefficients above .70. The factor intercorrelations were relatively high (correlations ranged from *r* = .53 to *r* = .91; average *r* =. 75), although they were lower compared to those reported by Li and Harmer (1996). Separate gender analysis showed that the GEQ had different factorial structures in the male and female samples, but none of these structures fitted the data well. In another study using confirmatory factor analysis, Sullivan, Short and Cramer (2002) found that the four-factor structure of the GEQ did not hold up when tested with athletes from co-acting teams.

In response to the literature (e.g., Schutz et al., 1994; Sulivan et al., 2002) criticising the psychometric properties of the GEQ, Brawley and Carron (2003) argued that some teams may not exhibit every factor of cohesiveness measured by the questionnaire. Furthermore, they suggested that researchers should make *a priori* hypotheses regarding the number of cohesion factors that should be identified in the groups under investigation. However, from a practical perspective this is not often possible as it would require a significant insight of each group under investigation. Furthermore, in order that cohesion research achieves such an *a priori* insight, the GEQ should be previously tested with samples with diverse characteristics.

Li and Harmer (1996) and Schutz et al. (1994) have also called for further testing of the GEQ using diverse samples. The purpose of the present study is to test the GEQ in a different culture (i.e., Greek). Studies in the sport psychology literature examining the psychometric properties of established English-language questionnaires in different cultures and languages have increased over the last decade (e.g., Isogai et al., 2001; Li, Harmer, Chi and Vongjaturapat, 1996). This is because cultural or linguistic characteristics may affect the applicability or interpretation of questionnaire items in different cultures. Besides testing the psychometric properties of the GEQ in a different culture, the present study aimed to examine the GEQ's factorial structure across the two gender groups and among elite and non-elite athletes. Schutz et al. (1994) argued that potential gender differences in the psychometric properties of the GEQ have been ignored in the literature. In fact, their analyses showed different factorial structures for males and females. However, since none of these structures had adequate model fit, any gender differences should be viewed with caution and need to be replicated. Lastly, differences in the factorial structure of the GEQ among elite and non-elite samples should be investigated. Carron (1982) argued that high level teams will give more emphasis on task cohesion than social cohesion. However, whether such differential emphasis has implications for the factorial structure of the GEQ across different competitive levels has yet to be tested. Furthermore, most studies in the group cohesion literature have examined amateur or semi-professional athletes. It would be important to examine whether the GEQ can be used with elite professional athletes. The findings of this study could be of value to coaches or sport psychologists who are interested to measure cohesion in their teams.

Method

Participants

The participants were volleyball players from the Greek Premier League and the Athens League (a regional league). Questionnaires were administered to 586 players, ranging in age from 18 to 34 years (*M*= 22.1; *SD*=4.7), with competitive experience varying from 1 to 15 years (*M*=4.7, *SD*=3.2). There were 130 males and 152 females from the premier league, and 151 males and 153 females from the Athens League.

Measure

Cohesion was assessed with the Group Environment Questionnaire (GEQ; Carron, et al., 1985) adapted to the Greek language by Aggelonidis (1995). The GEQ was derived from Carron et al.'s conceptual model of cohesion that views cohesion as a multidimensional construct. The questionnaire consists of 18 items scored on a 9-point scale ranging from 1 ('strongly disagree') to 9 ('strongly agree'). The questionnaire measures four dimensions of cohesion: ATG-T (e.g., 'I am not happy with the amount of playing time I get'; reverse item), ATG-S (e.g., 'Some of my best friends are on this team'), (GI-T; e.g., 'Our team is united in trying to reach its goals for performance'), and GI-S; ('Our team would like to spend time together in the off-season'). The ATG-T and GI-S subscales consist of 4 items each, whereas the ATG-S and GI-T subscales have 5 items each.

Procedure

The Greek GEQ was administered in a training session after obtaining written consent from the volleyball players and their coaches. The questionnaire was administered on four occasions, that is, at the first gathering of the new season ('pre-season'), at the beginning of the competitive season ('early season'), at the end of the first round ('mid-season'), and finally at the end of the second round ('end of season').

Data Analysis

Descriptive statistics, correlations and Cronbach alpha coefficients were first calculated. Then, the changes in the four cohesion dimensions across time were investigated using Linear Growth Modelling. Subsequently, a number of competing factorial structures were tested using confirmatory factor analysis. Finally, the invariance of the factorial structures across competitive level and gender was examined.

Results

Descriptive Statistics, Correlations and Internal Reliabilities

Table 1 shows the descriptive statistics and Cronbach alpha coefficients for the four cohesion subscales across the four time points. The means for the two social cohesion subscales remained stable across time with the ATG scores being consistently higher than the GI ones. In contrast, the means for the two task cohesion subscales were very similar, but both progressively declined across the four time points. All scores were above the midpoint (i.e., 5) of the rating scales. The standard skewness and kurtosis scores for all cohesion dimensions across all time points were small indicating that the data were normally distributed. The alpha coefficients for all subscales across all time measures were above .90 with some approaching unity.

Please insert Table 1 about here

To examine the statistical significance of the changes in the mean scores, latent growth modelling (LGM) was used for each of the four subscales. Due to the very large number of missing cases at the end of the season (i.e., the last time measure; see Table 1), the analysis was carried out only across the first three time points. LGM draws on the strengths of structural equation modelling and attempts to estimate the latent factors that represent the growth trajectories thought to underline the observed repeated measures of a variable (Curran and Hussong, 2002). These growth factors are the initial level (intercept) of the variable under study and the rate of change (slope) of its scores over time. With LGM it is possible to model the individual variability in the intercepts and slopes, thus estimating both the fixed (or group level) and random (or individual level) components associated with developmental trajectories. Meredith and Tisak (1990) showed that traditional repeated measures ANOVA designs are special cases of LGM, however, the latter provide more flexibility in the measurement of change (for a discussion of the advantages of LGM over repeated measures ANOVA, see Duncan, Duncan, Strycker, Li and Alpert, 1999). Because we analysed three time points, only linear growth models could be tested. Each cohesion factor was assessed using one observed score for each time point which represented the average of all items that composed the factor. We did not model the individual items because of the relatively small number of participants ($n = 197$) who completed the questionnaire across all three time points, and because confirmatory factor analyses (see below) showed that the convergent validity of the questionnaire was good.

For each cohesion dimension, a fixed effects or unconditional model was tested first in which the variance in the intercepts and slope factors was not modelled as a function of any explanatory variables. However, in cases where the variance in the intercept or slope was significant, a conditional model was tested in which the variance was modelled as a function of two explanatory variables, competitive level and gender. When specifying a LGM, the paths from the intercept to the repeated measures are fixed to 1. In linear LGM, the paths from the slope to the repeated measures are evenly spaced (i.e., 0, 1, 2, 3, etc..). Furthermore, the intercept and slope factors are allowed to covary. This covariance indicates whether the initial levels of a cohesion dimension are related to rates of subsequent change. For example, a negative covariance would indicate that individuals with higher initial levels of a given cohesion dimension displayed lower rates of change. An example of a linear conditional LGM model is presented in Figure 1.

Please insert Figure 1 about here

With regard to the two task cohesion scales (ATG-T and GI-T), the LGM results indicated that the mean of the intercept (i.e., initial level) and slope factors (i.e., change over time) were significant. The LGM results also indicated that there was significant variance in the intercept factor but not in the slope. This suggested significant individual differences in the starting point of ATG-T and GI-T but not in the trajectories of change of the two dimensions. The correlations between the intercept and the slope factors were non-significant. In an attempt to account for the individual differences in the starting point of ATG-T and GI-T, a conditional model was specified, separately for the two cohesion dimensions, with gender and competitive level as predictors. The results of the conditional model showed that the elite volleyball players had higher initial levels of ATG-T compared to the regional level ones. The gender differences were not significant. With regard to GI-T, none of the two individual difference factors could account for the significant variance in the intercept.

In relation to the two social cohesion factors (ATG-S and GI-S), the results of the unconditional model indicated a significant mean for the intercept but not for the slope, suggesting no significant changes in the scores over time. The correlations between the intercept and the slope were negative indicating that volleyball players with higher ATG-S and GI-S scores displayed lower rates of change across time. Furthermore, the LGM results indicated that there was significant variance in both the intercept and the slope of ATG-S, and in the slope of GI-S. A conditional LGM model showed that regional volleyball players had higher initial levels (intercept) of ATG-S and GI-S compared to elite players. No gender differences emerged for ATG-S, but with regard to GI-S, males had a higher slope (i.e., displayed steeper increases) than females. In all LGM models, the fit indices (see below for a more detailed description) were excellent.

Factorial Validity

To examine the factorial validity of the Greek GEQ, we carried out confirmatory factor analysis which tested a number of competing factor structures. In our analyses we tested the same factorial structures that were tested by Li and Harmer (1996). We first examined four different first-order factor models (see Figure 2): An one-factor model representing an overall cohesion factor (M1), a two-factor model representing the ATG and GI factors (M2), a two-factor model representing task and social cohesion (M3), and a four-factor model representing the four cohesion factors (M4) hypothesised by Carron et al. (1985). Due to the very large number of missing cases at the end of the season, the analyses were performed using the first three time points.

Please insert Figure 2 about here

The model fit was evaluated with the χ^2 statistic, the Non-Normed Fit Index (NNFI), the Comparative Fit Index (CFI), the Standardized Root Mean Square Residual (SRMR), and the Root Mean Square of Approximation (RMSEA). Due to the sensitivity of the χ^2 statistic to sample size, we based our interpretation of model fit mainly on the other fit indices using the cut-off criteria proposed by Hu and Bentler (1999). According to these criteria, a good model fit is shown by values close to .95 for NNFI and CFI, .07 for SRMR and .06 for RMSEA. Furthermore, for model comparison purposes we used the CAIC index instead of the chi-square difference test, because the latter, similar to the chi-square test, is sensitive to sample size (Biddle et al., 2000). There are no cut-off values for CAIC but, among competing models, lower values indicate better fit. In evaluating the appropriateness of a model we kept in mind Hair, Anderson, Tatham and Black’s (1998) caution to researchers to base their conclusions not only on model fit indices, but also on the appropriateness of the parameter estimates. For example, error variances which are negative or constrained to zero, and correlations or factor loadings which are very close to or exceed unity render problematic the interpretation of the model even hen its fit indices are excellent.

As can be seen in Table 2, the fit indices for each first-order factor model across the three time points were relatively similar.

Please insert Table 2 about here

The one-factor cohesion model (M1) had relatively good fit indices, although they were not quite as good as to meet Hu and Bentler’s (1999) cut-off criteria. All parameter estimates were in order. For illustration purposes the standardised loadings and uniqueness terms of M1 at pre-season are presented in Table 3.

Please insert Table 3 about here

The size of the loadings was relatively high (median = .83) providing support for the convergent validity of the questionnaire. Model 2 (M2) with two first-order AG and GI factors had nearly identical fit to M1, but the correlations between the two factors were very high (r = .98 for time 1 and 3 and r =.96 for time 2), casting doubt on the discriminant validity of the questionnaire. The two first-order task and social cohesion factor model (M3) had slightly better fit indices than M1 and M2 and a lower CAIC, but the correlations between the two factors were also very high (r = .96 for time 1 and 2, and r =.95 for time 3) rendering the interpretation of the model problematic. Lastly, the four factor cohesion model (M4) had similar fit to M3, however, the correlations among the four factors across the three time points were also very high ranging from .90 to 1.

Following the steps of Li and Harmer (1996), we also tested three second-order (hierarchical) factor structures. A hierarchical model is nested under its equivalent first-order model because it attempts to explain the correlations among the first-order factors in terms of fewer higher-order factors. When the fit of a higher order model is identical or very similar to the fit of the corresponding first-order model, the hierarchical model should be preferred because it is more parsimonious (Marsh, 1987). The first hierarchical model (M5) postulated that the four cohesion subscales could be subsumed under a general cohesion factor. The second hierarchical model (M6) postulated two second-order task and social cohesion factors underpinned by the respective ATG and GI first-order factors. Lastly, a third hierarchical model (M7) was tested based on Carron et al.'s (1985) model, which hypothesized two second-order factors, ATG and GI factors, underpinned by the respective task and social first-order factors. M5 had acceptable fit indices (with the exception of the RMSEA which was relatively high), but the second-order factor loadings were very high (they ranged from .95 to 1). M6 had also acceptable fit indices and one of the lowest CAIC scores compared to the other models. However, the correlation between the second-order task and social factors approached unity (r = .98 at time 1 and 2, and r =.97 at time 3), also rendering the interpretation of this model problematic. Lastly, M7 had similar fit indices to the other hierarchical models, however, the factor correlation between the ATG and GI factors was r = 1 on all three occasions[1].

Multisample analyses were also carried out to examine the invariance of all first- and second-order factor models across competitive level and gender. Due to the similar fit of each model across the three time points, the multisample analyses were carried out using the pre-season data only. We did not constrain the errors terms in any of the multisample analyses because the equality of error terms is a very restrictive hypothesis (see Byrne, 1994). The fit indices for the constrained models are presented in Table 4. These indices should be compared with the fit indices of the corresponding unconstrained models in Table 2. In all first-order factor models, we constrained the factor loadings and factor variances to be equal across groups. In addition, in M2, M3 and M4 we constrained the factor correlations to be equal. With regard to the second-order models M5, M6, and M7, we constrained the loadings of the first and second-order factors and the variances of the second-order factors to be equal across groups. The first-order factors do not have variances because they are dependent variables. In addition, in M6 and M7 we constrained the correlations between the second-order factors.

Please insert Table 4 about here

With regard to competitive level, when comparing each model in Table 4 with its corresponding unconstrained model (at pre-season) in Table 2, one can ascertain that the imposition of constraints did not result in noticeable changes in model fit. In fact, the relatively high RMSEA values were reduced to acceptable levels. Furthermore, the differences between the CFI fit indices of the constrained and unconstrained models were no more than .01. A difference in CFI values of .01 or below indicates that the constrained model should be preferred because it is not substantially worse compared to the unconstrained one (Cheung and Rensvold, 2002). However, in all seven models, the Lagrance Multiplier modification indices suggested partial model invariance (Byrne, 1994), because two of the constrained parameters were not invariant across the two competitive levels. These parameters were the factor loadings of two items on the ATG-T and GI-S factors respectively. Despite the two non invariant parameters, all other parameters (18-24 depending on the model) were invariant indicating that the factorial structure implied by each competing model (i.e., M1-M7) was largely the same across the two competitive levels.

With regard to gender, the multisample analyses showed that the constrained models had very similar fit compared to the corresponding unconstrained models. Furthermore, all parameter estimates were invariant across all seven models with the exception of the correlation between GI-S and ATG-T in M4 which was not invariant. Therefore, the results showed that in six of the seven tested models, the factorial structure of the Greek GEQ was the same in the two gender groups.

Discussion

The purpose of the present study was to examine the psychometric properties of the GEQ with a sample of Greek elite and regional level volleyball players. Data from three separate occasions were analysed and the results were very similar, thus giving us confidence to make some inferences regarding the Greek GEQ's psychometric properties.

Initial descriptive statistics analysis showed that the four subscales of the Greek GEQ (ATG-T, ATG-S, GI-T, and GI-S) were normally distributed since all the standard kurtosis and skewness scores were small. Similar findings were reported by Schutz et al. (1994). All mean scores were somewhat above the midpoint of the scale (i.e., 5) and the standard deviations were relatively small in comparison to the respective means. An analysis of the changes in the mean scores across three time points (i.e., pre-season, beginning of season, and mid-season) showed a decline in the task cohesion scores but no significant change in the social cohesion ones. The high score for the task cohesion scales at pre-season may reflect the fact that teams give high emphasis on team-related strategies and interactions during practice as they prepare for the upcoming competitive season. These strategies are consolidated by the start of the season and may not be emphasized as much later on in the season as during the pre-season, which might explain the slight drop in the task cohesion scores. However, across all time points the task cohesion scores were relatively high. Elite players had significantly higher ATG-T scores but significantly lower social cohesion scores (both ATG-S and GI-S) compared to the regional level players. This difference between the two competitive levels is not surprising given that elite teams are highly task-oriented groups. In such groups the emphasis is more on achieving the group objectives than on developing and maintaining social relationships (Carron, Brawley and Widmeyer, 1998). In contrast, it seems that in lower competitive levels athletes are mainly attracted by the social aspects of sport participation. It would be interesting for future research to examine whether such an attraction can predict over time athletes' commitment to their teams.

All subscales had very high internal reliability coefficients with Cronbach's alphas above .90. Whilst some psychometrics experts argue that psychological tests should be as internally consistent as possible (e.g., Nunnally and Bernstein, 1994), others argue that very high internally consistency indicates item redundancy since the test items are essentially paraphrases of each other (e.g., Cattell, 1973). In fact, high internal reliability scores increase as the correlations among the items increase (Kline, 2000). In the present study, the bivariate correlations among the questionnaire items were very high (ranging from $r = .75$ to $r = .90$), and this was reflected in the near perfect correlations among the latent factors.

Such high correlations proved to be problematic when examining the factorial validity of the Greek GEQ. Various competing factor structures were tested using confirmatory factor analysis. The results were very consistent across all three time points and showed that most of the competing models had acceptable model fit. This is partly due to the fact that all factor loadings were moderate to large indicating good convergent validity. For comparison purposes, the median of the loadings for M4 reported by Li and Harmer (1996) was .67, the median reported by Schutz et al. (1994) was .52, whereas in the present study it was .85. In relation to the first-order factor models tested, the four-factor model (M4) had the best fit indices and the lowest CAIC value. Li and Harmer (1996) also found that the four-factor model had the best fit indices among the first-order models they tested with a sample of US intercollegiate athletes. However, the correlations among the four factors in our study were very high and in some cases equalled 1. Such high correlations indicate that the Greek GEQ has low discriminant validity. Li and Harmer (1996) also reported high factor correlations among the GEQ subscales, although with an average of $r = .82$, these correlations were not as high as those in the present study. The fit of the two other first-order factor models (M2 and M3) postulating ATG/GI and task/social factors respectively was also relatively good, however, the factor intercorrelations were again very high. The one-factor cohesion model (M1) had the worse fit indices of the four first-order model, however, in absolute sense, the fit indices were close to the cut-off values recommended by Hu and Bentler (1999).

Following the example of Li and Harmer (1996), we also examined three hierarchical models since the high correlations among the first-order factors implied a hierarchical structure for the Greek GEQ. Furthermore, from a conceptual viewpoint, Carron et al.’s (1985) theoretical model has a hierarchical structure. All three hierarchical models had a good fit, but similar to the first-order models, the second-order factor correlations were very high, implying again poor discriminant validity. As a matter of interest, it should be noted that the correlation between the ATG and

GI second-order factors reported by Li and Harmer (1996) was also very high (*r* = .91). The good fit of most models in this study stands in contrast to the findings reported by Schutz et al. (1994) whose attempts to fit their data to first and second-order factor models were unsuccessful. This might be due to the fact that Schutz et al.'s (1994) study used high school athletes, whereas this study used predominantly adult professional and semi-professional athletes. In fact, the sample of this study is more similar to the college and adult athletes used in the initial validation studies of the GEQ (for a similar argument, see also Li and Harmer, 1996).

Multisample analyses were also conducted to examine the invariance of the first- and second-order factor models separately for elite and regional level athletes as well as for males and females. The results showed that the Greek GEQ was largely invariant across the different competitive level and gender groups. However, the lack of discriminant validity observed in the whole sample was also observed in the different sub-samples. Future studies need to examine the factorial invariance of the original GEQ across different gender and competitive level groups because it is not known whether the largely invariant factor structures found in this study would be also observed with the English language version of the questionnaire.

Taken together, these findings indicate that the Greek GEQ has acceptable fit indices and good convergent validity, however, its discriminant validity is poor. From a practical perspective, the results imply that coaches or sport psychologists employing the Greek GEQ could use an overall group cohesion score since the athletes cannot clearly differentiate among its subscales. In fact, the fit of the one-factor cohesion model (M1), although did not meet the cut-off criteria recommended by Hu and Bentler (1999), was relatively good. However, further testing of this questionnaire with Greek samples from different sports is needed because the psychometric evidence presented here is preliminary. Based on the existing evidence, one could argue that the GEQ has higher than anticipated factor correlations in both English (see Li and Harmer, 1996) and Greek languages and perhaps the items should be reworded to become more independent. An alternative explanation for the high factor correlations is that coaches might regard all aspects of cohesion as highly interwoven and, therefore, they might implement strategies to promote all of them in their teams. For the further psychometric validation of the Greek GEQ, future research should examine whether it relates to the same personal (e.g., satisfaction) and team variables (e.g., group performance) that the English GEQ does.

References

Aggelonidis, I. (1995) *Team Cohesion and Coach Leadership Style as Factors Affecting the Performance of Volleyball Teams*. Unpublished doctoral dissertation, University of Athens, Greece.

Biddle, S., Markland, D., Gilbourne, D., Chatzisarantis, N. and Sparkes, A. (2001) ‘Research Methods in Sport and Exercise Science: Quantitative and Qualitative Issues’, *Journal of Sports Sciences, 19,* 777-809.

Brawley, L.R. and Carron, A.V. (2003) ‘Caution in Interpreting Confirmatory Factor Analysis of the Group Environment Questionnaire: A Response to Sullivan, Short, and Cramer (2002)’, *Perceptual and Motor Skills, 97,* 315-318.

Byrne, B. M. (1994) *Structural Equation Modeling with Eqs and Eqs/Windows: Basic Concepts, Applications, and Programming.* Thousand Oaks, CA: Sage.

Carron, A.V. (1982) ‘Cohesiveness in Sport Groups: Interpretations and Considerations’, *Journal of Sport Psychology, 4,* 123-138.

Carron, A.V., Brawley, L.R. and Widmeyer, W.N. (1998) ‘The Measurement of Cohesiveness in Sport Groups’. In J.L. Duda (ed.), *Advances in Sport and Exercise Psychology Measurement,* pp. 213-226. Morgantown, WV: Fit Press.

Carron, A. V. and Hausenblas, H. A. (1998) *Group Dynamics in Sport* (2nd ed.). Morgantown, WV: Fitness Information Technology.

Carron, A.V., Widmeyer, W.N. and Brawley, L.R. (1985) ‘The Development of an Instrument to Assess Cohesion in Sport Teams: The Group Environment Questionnaire’, *Journal of Sport Psychology, 7,* 244-266.

Cattell, R. B. (1973) *Personality and Mood by Questionnaire*. San Francisco: Jossey-Bass.

Cheung, G., W. and Rensvold, R. B. (2002) ‘Evaluating Goodness-of-Fit Indexes for Testing Measurement Invariance’, *Structural Equation Modeling: A Multidisciplinary Journal, 9*, 233-255.

Curran, P.J. and Hussong, A.M. (2002) ‘Structural Equation Modeling of Repeated Measures Data’. In D. Moskowitz and S. Hershberger (eds.), *Modeling Intraindividual Variability with Repeated Measures Data: Methods and Applications*, pp. 59-86. New York: Erlbaum Associates.

Duncan, T. E., Duncan, S. C., Strycker, L. A., Li, F. and Alpert. A. (1999) *An Introduction to Latent Variable Growth Curve Modeling: Concepts, Issues, and Applications*. Mahwah, NJ: Lawrence Erlbaum Associates**.**

Hair, J.F., Anderson, R.E., Tatham, R.L. and Black, W.C. (1998) *Multivariate Data Analysis* (5th ed.). Upper Saddle River, NJ: Prentice Hall.

Hu, L. and Bentler, P. M. (1999) ‘Cutoff Criteria for Fit Indexes in Covariance Structure Analysis: Conventional Criteria versus New Alternatives’, *Structural Equation Modeling, 6*, 1-55.

Isogai, H., Brewer, B.W., Cornelius, A.E., Komiya, S., Tokunaga, M. and Tokushima, S. (2001) ‘Cross-Cultural Validation of the Social Physique Anxiety Scale’, *International Journal of Sport Psychology, 32,* 76-87.

Kline, P. (2000) *The Handbook of Psychological Testing* (2nd ed.). London: Routledge.

Li, F. and Harmer, P. (1996) ‘Confirmatory Factor Analysis of the Group Environment Questionnaire with an Intercollegiate Sample’, *Journal of Sport and Exercise Psychology, 18,* 49-63.

Li, F., Harmer, P., Chi, L. and Vongjaturapat, N. (1996) ‘Cross-Cultural Validation of the Task and Ego Orientation in Sport Questionnaire’, *Journal of Sport and Exercise Psychology, 18,* 392-407.

Marsh HW. (1987) ‘The Hierarchical Structure of Self-Concept and the Application of Hierarchical Confirmatory Factor Analysis’, *Journal of Educational Measurement, 24*: 17-39.

Meredith, W. and Tisak, J. (1990) ‘Latent Curve Analysis’, *Psychometrika, 55,* 107-122.

Nunnally, J.C. and Bernstein, I.H. (1994) *Psychometric Theory* (3rd ed.). New York: McGraw-Hill.

Schutz, R.W., Eom, H.J., Smoll, F.L. and Smith, R.E. (1994) ‘Examination of the Factorial Validity of the Group Environment Questionnaire’, *Research-Quarterly for Exercise and Sport, 65*, 226-236.

Sullivan, R.A., Short, S.E. and Cramer, K.M. (2002) ‘Confirmatory Factor Analysis of the Group Environment Questionnaire with Coacting Teams’, *Perceptual and Motor Skills, 94,* 341-347.

Widmeyer, W.N., Brawley, L.R. and Carron, A.V. (1985) *The Measurement of Cohesion in Sport Teams: The Group Environment Questionnaire.* London, ON: Sport Dynamics.

Footnote

1. Following a suggestion by an anonymous reviewer, a third-order factor structure was also tested for M6 and M7, with a third-order “cohesion” factor underpinned by two second-order factors. The fit of the third-order factor models was almost identical to the fit of the corresponding second-order models, thus supporting a third-order factor structure. However, such a structure is not discussed by Carron et al. (1985). Furthermore, identification problems were manifested in the solution (i.e., error variances constrained to zero), probably due to the small number of higher-order factors, therefore the fit of the third-order factor models should be viewed with caution.

Table 1

Descriptive Statistics and Internal Reliability Cronbach Alpha Coefficients for the Four Cohesion Subscales across the Four Time Points.

	M	*SD*	Standard Skewness	Standard Kurtosis	*N*	α
Pre-season						
ATG-S	6.67	1.60	-0.65	-0.15	452	.92
ATG-T	6.80	1.63	-0.68	0.13	439	.91
GI-S	5.27	1.87	-0.10	-0.67	449	.91
GI-T	6.60	1.55	-0.55	-0.13	449	.94
Early season						
ATG-S	6.59	1.65	-0.69	0.09	411	.95
ATG-T	6.49	1.73	-0.45	-0.51	410	.92
GI-S	5.36	1.87	-0.10	-0.66	408	.93
GI-T	6.43	1.61	-0.41	-0.46	407	.95
Mid-season						
ATG-S	6.48	1.65	-0.75	0.45	364	.96
ATG-T	5.80	1.90	-0.23	-0.54	362	.92
GI-S	5.31	1.96	-0.05	-0.86	362	.95
GI-T	5.88	1.75	-0.47	-0.08	361	.95
End of season						
ATG-S	6.72	1.65	-0.81	0.07	149	.98
ATG-T	5.71	1.90	-0.13	-0.71	149	.95
GI-S	5.61	1.99	-0.23	-0.72	150	.98
GI-T	5.81	1.71	-0.22	-0.38	149	.97

Note: ATG-S Individual Attraction to the Group-Social; ATG-T= Individual Attraction to the Group-Task; GI-S= Group Integration-Social; GI-T= Group Integration Task.

Table 2

Fit Indices of the Competing Models Across the First Three Time Measures.

	χ^2	df	NNFI	CFI	SRMR	RMSEA	CAIC
Pre-season							
Model 1	891.48	135	.92	.93	.03	.10	-111.29
Model 2	767.28	134	.93	.94	.03	.09	-220.74
Model 3	754.042	134	.94	.94	.03	.09	-233.98
Model 4	584.99	129	.95	.96	.03	.08	-366.17
Model 5	622.39	131	.95	.96	.03	.08	-343.51
Model 6	577.13	130	.95	.96	.03	.08	-381.40
Model 7	621.77	130	.95	.96	.03	.08	-366.76
Early season							
Model 1	973.63	135	.93	.94	.03	.10	-21.77
Model 2	915.040	134	.93	.94	.03	.10	-72.98
Model 3	792.20	134	.94	.95	.03	.09	-195.82
Model 4	598.29	129	.96	.96	.02	.08	-352.87
Model 5	635.16	131	.96	.96	.02	.08	-330.74
Model 6	609.81	130	.96	.96	.02	.08	-348.72
Model 7	635.10	130	.96	.96	.02	.08	-323.43
Mid-season							
Model 1	1162.50	135	.92	.93	.03	.11	167.33
Model 2	1070.94	134	.92	.93	.03	.11	83.14
Model 3	799.32	134	.95	.95	.02	.09	-188.48
Model 4	606.67	129	.96	.97	.02	.08	-344.27
Model 5	422.17	131	.95	.96	.03	.09	-281.51

Model 6	616.44	130	.96	.97	.02	.08	-341.87
Model 7	679.01	130	.95	.96	.03	.09	-279.30

Note: All chi-square values are significant at $p < .001$

Table 3

Standardized Loadings and Uniqueness Terms of Model 1 at Pre-Season.

Items	Loadings	Uniqueness
ATG-S1	0.79	0.62
ATG-T1	0.78	0.63
ATG-S2	0.88	0.49
ATG-T2	0.82	0.57
ATG-S3	0.72	0.69
ATG-T3	0.84	0.54
ATG-S4	0.83	0.56
ATG-T4	0.88	0.47
ATG-S5	0.88	0.48
GI-T1	0.90	0.43
GI-S1	0.78	0.63
GI-T2	0.85	0.53
GI-S2	0.74	0.67
GI-T3	0.81	0.59
GI-S3	0.81	0.59
GI-T4	0.88	0.47
GI-S4	0.83	0.56
GI-T5	0.85	0.52

Note: ATG-S Individual Attraction to the Group-Social; ATG-T= Individual Attraction to the Group-Task; GI-S= Group Integration-Social; GI-T= Group Integration Task.

The parameter estimates at the start of the season and at mid-season were very similar.

Table 4

Fit Indices of the Invariance Testing Of M1-M7 for Elite and Regional Volleyball Players (and for Males and Females in Brackets) at Pre-Season.

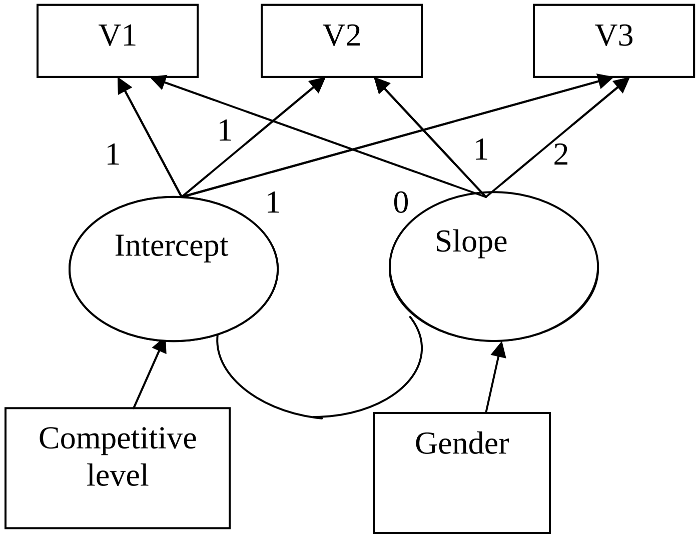
	Constraints imposed	χ^2	df	NNFI	CFI	SRMR	RMSEA	CAIC
Model 1	Equal first-order factor loadings and	1124.02	288	.92	.93	.05	.07	-999.49
	variances	(1164.20)	(288)	(.92)	(.92)	(.06)	(.07)	(-959.31)
Model 2	As in M1 and with equal first-order	997.51	287	.93	.94	.05	.07	-1118.63
	factor correlations	(1028.76)	(287)	(.93)	(.93)	(.06)	(.07)	(-1087.38)
Model 3	As in M2	989.91	287	.93	.94	.05	.07	-1126.24
		(1032.10)	(287)	(.93)	(.93)	(.06)	(.07)	(-1084.05)
Model 4	As in M2	757.29	282	.95	.96	.03	.06	-1311.19
		(794.48)	(282)	(.95)	(.95)	(.06)	(.06)	(-1284.80)
Model 5	Equal first and second-order factor	811.49	280	.95	.95	.05	.06	-1253.03
	loadings and equal second-order	(848.72)	(280)	(.95)	(.95)	(.06)	(.06)	(-1215.81)

	factor variances							
Model 6	As in M5 and with equal second-	803.55	279	.95	.95	.05	.06	-1253.61
	order factor correlation	(841.70)	(279)	(.95)	(.95)	(.06)	(.06)	(-1215.46)
Model 7	As in M6	796.41	279	.95	.95	.05	.06	-1260.74
		(832.12)	(279)	(.95)	(.95)	(.06)	(.06)	(-1225.04)

Note: All chi-square values are significant at $p < .001$

Figure 1

An Example of a Conditional Linear Latent Growth Model

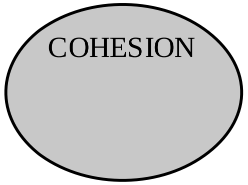


Note: V1, V2, and V3 represent the average scores of a cohesion factor for the first three time points respectively. All variables and factors in the model are regressed on a constant which is not shown here for presentation simplicity reasons. The unconditional model does not include competitive level and gender.

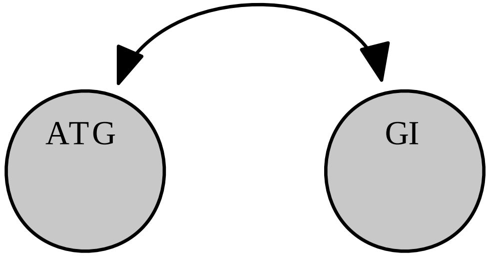
Figure 2

Competing Confirmatory Factor Analysis Models

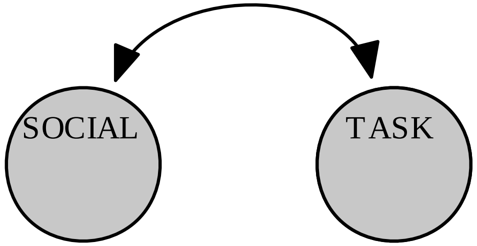
Model 1: A first-order one factor Cohesion model



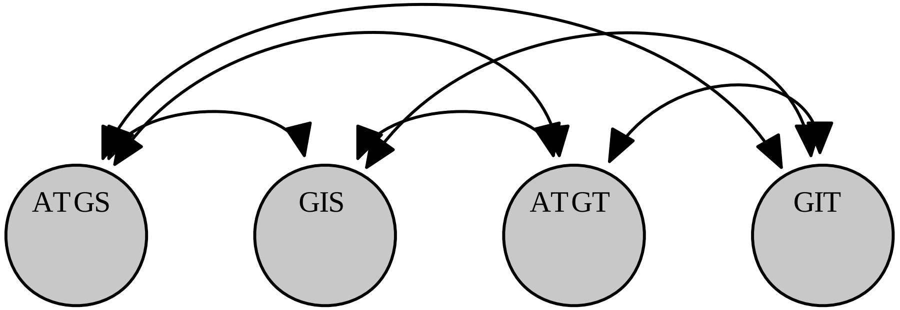
Model 2: A first-order two factor Attraction to the Group and Group Integration model



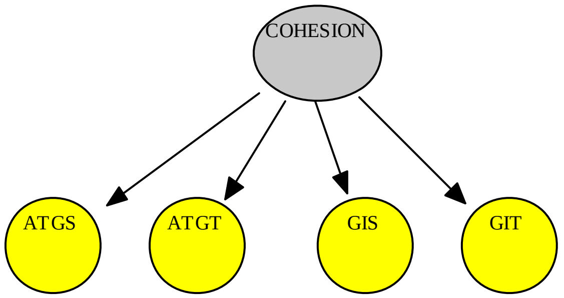
Model 3: A first-order two factor Task and Social cohesion model



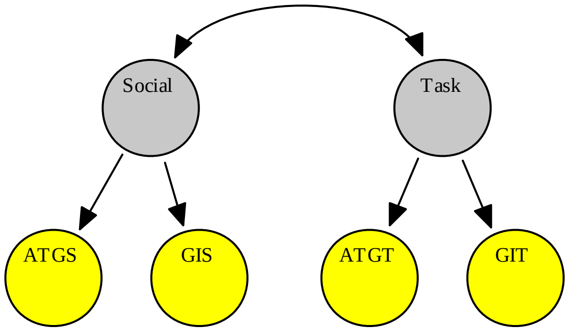
Model 4: A first-order four factor model



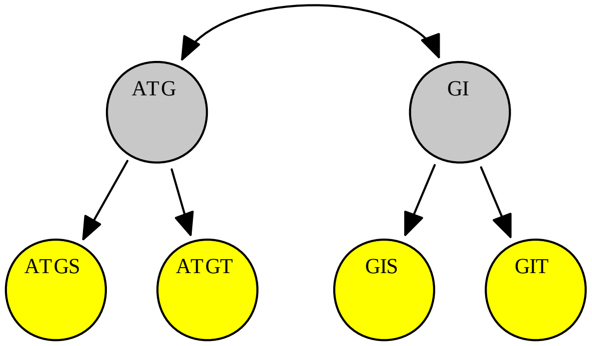
Model 5: A second-order one factor Cohesion model



Model 6: A second-order two factor Task and Social cohesion model.



Model 7: A second-order two factor Attraction to the Group and Group Integration model.



Note: The observed indicators and the residual terms have been omitted from the graphs for presentation simplicity reasons.