Running Head: LONG VERSIONS OF THE CART-Q

Initial evidence for the criterion-related and structural validity of the long versions of the direct and meta-perspectives of the Coach-Athlete Relationship Questionnaire

Daniel J. A. Rhind – Brunel University

and

Sophia Jowett – Loughborough University

Paper submitted for publication in *The European Journal of Sport Science*

Date of submission: November 4th, 2008

Date of final submission: January 6th, 2010

Correspondence concerning this article should be addressed to Daniel J. A. Rhind, School of Sport and Education, Brunel University, Uxbridge, UB8 3PH

Tel.: +44 (0) 1895 266860

Email: Daniel.Rhind@Brunel.ac.uk

Abstract

The present study aimed to develop and initially validate a longer version of the Coach-Athlete Relationship Questionnaire (CART-Q), both its direct perspective (Jowett & Ntoumanis, 2004) and meta-perspective (Jowett, in press, 2007b). In Study 1, Instruments (e.g., questionnaires, scales and inventories) that have been used to assess relationship quality in the broader psychological literature were examined and items potentially relevant to the coach-athlete relationship were identified. The content validity of the identified items was then assessed using expert panels. A final questionnaire was subsequently prepared and administered to 693 participants (310 coaches and 383 athletes). Confirmatory Factor Analysis was employed to assess the multidimensional nature of the questionnaire based on the 3 Cs (i.e. closeness, commitment and complementarity) model of the coach-athlete relationship. The findings indicated that the direct and meta-perspective items of the long versions of the CART-Q approached an adequate data fit. Moreover, evidence for the internal consistency and criterion validity of the new instruments was also obtained. In Study 2 the newly developed measure was administered to an independent sample of 251 (145 athletes and 106 coaches). Further statistical support was gained for the factorial validity and reliability of the longer version of the CART-Q.

Keywords: coach-athlete relationship quality, validation, performance, satisfaction

Initial evidence for the criterion-related and structural validity of the long versions of the direct and meta-perspectives of the Coach-Athlete Relationship Questionnaire

The last decade has witnessed a significant increase in the amount of research being conducted regarding the coach-athlete relationship (Jowett & Wylleman, 2006). Research in this domain has been fuelled by the suggestion that positive outcomes in sport are associated with quality relationships between coaches and athletes (Petitpas, 2002). Furthermore, elite award-winning coaches have been found to place a tremendous emphasis on having quality coach-athlete relationships for both sporting success and the athlete's personal development (Gould, Collins, Louer & Chung, 2007). As a result, there is a need to investigate the nature and role played by this relationship and to establish how it should best be conceptualized, measured, and ultimately optimized.

The coach-athlete relationship has been defined as "…the situation in which coaches' and athletes' emotions, thoughts and behaviors are mutually and causally inter-connected" (Jowett & Ntoumanis, 2004, p. 245). This definition was developed based on research conducted in the area of close relationships and benefits from emphasizing the bi-directional nature of such relationships as well as incorporating the affective (emotions), cognitive (thoughts) and behavioral aspects (Cf., Kelley et al., 1983).

Jowett (2005, 2007a) proposed the 3+1C (i.e., closeness, commitment, complementarity and co-orientation) conceptualization of the coach-athlete relationship. The '3Cs' of this model represents the constructs of closeness (cf., Rubin, 1970), commitment (cf., Rusbult, Martz, & Agnew, 1998) and complementarity (cf., Kiesler, 1997). Closeness refers to the affective meanings that the coach and athlete assign to their relationship (e.g., liking, trust, and respect). Commitment relates to the members' intentions to maintain the relationship at present and in the future. Complementarity concerns the members' co-operative and corresponding behaviors of affiliation and hence

an athlete's friendly and responsive behavior during training attracts friendly and responsive behavior from their coach. The '+1Cs' of Jowett's (2005, 2007a) model is labeled as co-orientation which highlights the degree to which the athlete's and the coach's perceptions are interconnected and therefore reflect their common ground. Co-orientation is comprised of the direct perspective (how one person feels, thinks and behaves towards the other; e.g., 'I like my coach/athlete') and the meta perspective (how the athlete/coach believes their coach/athlete feels, thinks and behaves; e.g., 'My coach/athlete likes me'). A series of qualitative studies have been conducted to explore the nature of the coach-athlete relationship by assessing the existence and content of the 3 + 1Cs (e.g., Jowett, 2003; Jowett & Cockerill, 2003; Jowett & Meek, 2000). For example, the study by Jowett and Meek (2000) involved four coach-athlete dyads who were also married. In-depth interviews were conducted based on the interpersonal constructs of closeness, complementarity and co-orientation. A content analysis of the data revealed that the close relationship facilitated co-oriented views in terms of relevant tasks and goals which in turn contributed to co-operative acts on the training ground. Jowett and Cockerill (2003) interviewed 12 athletes who had all won an Olympic medal between 1968 and 1988. Themes relating to closeness (e.g., trust and respect), co-orientation (e.g., common goals) and complementarity (in terms of roles and tasks) were found to underline relationship quality as experienced by some of the most successful athletes in the world.

Based upon the data gained through the qualitative research studies, the 11-item direct perspective of the Coach-Athlete Relationship Questionnaire (CART-Q direct; Jowett & Ntoumanis, 2004) was developed. Jowett and Ntoumanis (2004) conducted two validation studies; the first study employed Principal Components Analysis (PCA) to uncover the underlying structure of the instrument, whilst the second study employed Confirmatory Factor Analysis (CFA) to confirm its underlying factor structure. The latter study demonstrated that both the three first order factor model (i.e., Closeness,

Commitment, and Complementarity; 3 Cs) and a higher order model which was hypothesized to represent the overall quality of the coach-athlete relationship and contained the 3 Cs, have the same satisfactory model fit indices (Comparative Fit Index (CFI) = 0.96, Non-Normed Fit Index (NNFI) = .94, Standardized Root Mean Square Residual (SRMR) = .05 and Root Mean Square Error of Approximation (RMSEA) = .09). Overall, Jowett and Ntoumanis (2004) provided initial evidence for the content, factorial and criterion-related validity as well as the internal consistency of the 11-item direct perspective of the CART-Q.

A meta perspective version of the 11-item CART-Q direct has also been developed in order to facilitate the assessment of the construct of co-orientation (see Jowett, 2005, 2007a). Thus, for each of the direct items, a meta item was created which allowed statements to be phrased in a way that considers how a coach or athlete believes the other member feels or thinks. The 11-item meta-perspective CART-Q has demonstrated corresponding factor structure with that of the direct CART-Q. Jowett (2007b) has found that the latent structure (e.g., a model which theorizes the relationships which exist between the variables and factors within the data) comprised of three first order factors (meta closeness, meta commitment and meta complementarity) had an excellent fit to the data (CFI = 1.00, SRMR = .04 and RMSEA = .00). A second model which also contained a higher order factor which represented a general meta relationship factor also had acceptable fit indices (CFI = 0.95, SRMR = .04 and RMSEA = .07). Subsequent empirical research has lent further support to the validity of the CART-Q direct and meta versions (see e.g., Jowett & Clark-Carter, 2006).

The CART-Qs have proved to be useful instruments for researchers interested in assessing the nature, content or quality of the coach-athlete relationship. In particular, the relatively short length of the questionnaires ensure that they can be administered expediently and as part of a battery of questionnaires within a research study (e.g., Jowett

& Chaundy, 2004; Jowett & Clark-Carter, 2006). However, it is possible that researchers and practitioners require detail not necessarily captured in the current short versions of the CART-Q. Thus, the option to select from either a short or a long version based on the requirements dictated by a researcher's and practitioner's specific work, suggested the need for the development and validation of longer versions of the CART-Q to fill the identified gap in relation to the available measures regarding the coach-athlete relationship (Wylleman, 2000).

Gill, Dzewaldowski, and Deeder (1988) argued for the construction of multi-dimensional instruments within sport psychology which are based on theory, item and reliability analysis, factor analysis, tests of convergent and divergent validity, validation in relation to external criteria and application in research and practice. In the sport psychology domain, the concepts of interest are generally hypothetical and therefore evidence of their construct validity needs to be demonstrated. Hence, Duda (2001) called for all sport-specific instruments to be evaluated within a construct validity framework.

Whenever one is developing a measure of a psychological construct, two crucial concepts to consider are validity and reliability. The American Psychological Association (1999) has argued for a unified validity which is comprised of face/content validity, criterion-related validity, and construct validity. Content validity relates to the "…match between items or tasks in the measure and the content domain to which generalization is sought" (Hoyt, Warbasse & Chu, 2006, p. 774). Moreover, Hoyt and colleagues (2006) have explained that criterion-related validity includes "…correlations with future performance (predictive validity) and correlations with theoretically-related constructs assessed at the same point in time" (concurrent validity) (p. 776). Finally, construct validity relates to "…an ongoing theory-guided enquiry into systematic determinants of test scores (often called the test's factor structure or internal structure" (Hoyt et al., 2006, p. 778). In a seminal chapter, Messick (1989) argued for multiple standards for the assessment of

construct validity such as relevant content (which is based on sound theory and rationale), internally consistent items, external correlations with related measures, generalisability across different samples and time, as well as being explicit in terms of social consequences (e.g., racial bias). Construct validity is also indicated by convergent validity (e.g., whether a measure is associated to other measures to which it should theoretically be linked with) and discriminant validity (e.g., whether a measure is not associated with other measures to which it should, in theory, not be linked with).

Reliability, or internal consistency, concerns how well items which are developed to measure a single construct correlate with each other. It therefore assesses whether the items are measuring the same construct (American Psychological Association, 1999). Thus, any newly developed or extended psychological instrument needs to show evidence of both validity and reliability for it to be deemed a sound scientific measure.

The focus of the present research is to develop longer versions of the direct and meta CART-Qs. The need for this expansion is based on a number of important theoretical, applied, and research-related factors. In theoretical terms, expanding the measures of closeness, commitment, and complementarity will broaden the scope of the relevant issues that are assessed and addressed within each subscale. On an applied level, a longer instrument is more likely to provide a more comprehensive and detailed assessment of this dyadic relationship. This information could be used for the development of interventions and norm data can also be developed which would further help to identify areas of strength or weakness within the coach-athlete relationship (e.g., high or low levels of commitment). This should help narrow the gap which has been identified between existing theoretical knowledge and its usefulness for coaches and athletes (Coppel, 1995). It will also meet the need for more research that has more practical use for coaches and athletes (Poczwardowski, Sherman, & Henschen, 1998) whilst addressing the need for the delineation of valid instrumentation in this domain as identified by Wylleman (2000). In

research terms, it would be possible for researchers who are interested in investigating specific aspects of the relationship (e.g., closeness, commitment, and/or complementarity), to use the longer versions as stand-alone measures.

It is the purpose of the present research to develop and initially validate a longer version of the direct and meta perspective of the CART-Questionnaires. A number of authors have proposed procedures for use when developing scales. The procedure used in developing these scales follow DeVellis' (2003) recommended eight stages: (1) determine clearly what you want to measure, (2) generate an item pool, (3) determine the format of the measure, (4) have experts review the initial item pool, (5) consider inclusion of validation items, (6) administer items to a development sample, (7) evaluate the items and (8) optimize the scale length. Two studies were conducted. Study 1 has three phases: Phases 1 and 2 cover stages 1- 4 and its aim was to demonstrate the content validity of the items to be included in the expanded questionnaires whilst Phase 3 covers stages 5 - 8 and its aim was to establish the criterion and construct validity of the longer scales. Study 2 was conducted in order to further investigate the factorial validity and reliability of the expanded instruments with an independent sample.

Phase 1: Item Generation

Computerized databases, key books, the internet, and the reference lists of relevant resources (e.g., book chapters, journal articles) were reviewed to identify potentially relevant measures developed to assess aspects of relationship quality. This included all forms of relationship and not just those in the sporting context. An article was retrieved if it appeared to relate to the measurement of either Closeness, Commitment or Complementarity, or a related concept, within human relationships. This ensured that research from a wide range of different relationships was included such as parent-child, doctor-patient and romantic partners. As a result of this process, 14 measures of closeness were retrieved, 11 measures of commitment and 9 measures of complementarity. All of the

items within these measures were then pooled to create three sets of relevant items (closeness = 290 items, commitment = 291 items, and complementarity = 278 items). The identification of scales was completed by the first author who has expertise in conducting such searches and experience of undertaking relationship research.

Each of the 859 items was then carefully considered by the authors to ensure that it was relevant and representative of one of the clearly defined constructs (i.e., one of the 3 Cs). This helped to reduce the risk of introducing error which would in turn negatively impact the strength of inter-item correlations and hence go against the objectives of scale development (Quintana & Minami, 2006). All efforts were made to ensure that all items were clear, concise, distinct, comprehensible, and reflected the construct of interest (Anastasi, 1988).

Each item was either included (with modifications made where appropriate) or excluded from further analysis. Items were excluded for a number of reasons. For instance, some items duplicated other included items. Other items addressed issues which were irrelevant to the coach-athlete relationship (e.g., related to sex) or contained words which were viewed as having the potential to cause comprehension problems (e.g., "My partner is antagonistic"). A total of 28 items for Closeness, 29 items for Commitment, and 38 items for Complementarity were retained following this process. All of the original 11 items in the CART-Q were retained.

Three separate documents were created following the procedures outlined by DeVellis (2003). Each contained items that assessed either closeness, commitment or complementarity. Four different versions of each of these three documents were developed to represent (a) athlete direct perspective, (b) athlete meta-perspective, (c) coach direct perspective, and (d) coach meta-perspective. In the direct versions the items focused on how the respondent felt, thought or behaved towards the other member of the dyad (e.g., "I like my coach/athlete"). The meta versions contained the same basic items but with them

being reworded such that they assessed how a respondent believed the other member of their dyad felt, thought or behaved (e.g., “My coach/athlete likes me”). Therefore a total of 12 documents were created.

Phase 2: Content validity

The aim of the second phase was to assess the content validity of the items identified in Phase 1. Panels of experts assessed the pool of items the authors had identified as potentially relevant in Phase 1. A total of 12 expert panels were employed (one for each of the 12 questionnaires developed in Phase 1) which were comprised of 44 experts. Experts consisted of 20 sport psychology consultants and/or academics, 12 sport psychology PhD students, 6 coaches and 6 athletes. Each expert was given a pack that contained instructions and the set of items that either dealt with the direct or meta-perspective of closeness, commitment or complementarity. This ensured that all items were reviewed and scrutinized by several members of a panel. Each expert panel contained at least one subject matter expert (e.g., sport psychology academic and/or consultant), a sport psychology Ph.D. student, and either a coach or an athlete (depending upon the specific version of the questionnaire being assessed). Thus each judge reviewed either the coach direct, coach meta, athlete direct or athlete meta version of the items designed to measure closeness, commitment or complementarity.

Experts independently read through the set of items and indicated whether they thought each of them were “representative’ (i.e. does it represent the construct of interest”, “Relevant” (i.e., does it reflect the definition of Closeness/Commitment/Complementarity provided), “Clear” (i.e., is it easily understood), and “Specific” (i.e., is it focused and not too general or ambiguous). Participants indicated their opinions by circling “Yes”, “No” or “Unsure” for whether each item was representative, relevant, clear or specific. A “Comments” section was also included under each item to enable the participant to explain

their responses and to suggest any alterations. Finally, panel members were asked about the appropriateness of the pitch, flow, instructions used, and presentation of the questionnaire.

The feedback from the expert panels was then reviewed. An item was retained if it was viewed as relevant, representative, clear and specific by at least 9 out of the 12 panel members (3/4) and if no significant issues had been highlighted by any respondent. This procedure was adopted as it has been employed in previous related research by Jowett and Ntoumanis (2004) in the development of the short version of the CART-Q. This resulted in three item pools being generated containing 21 items (closeness), 20 items (commitment) and 23 items (complementarity). Evidence for the content validity of all of these 64 items was found.

Phase 3: Construct and criterion validity

The aim of this phase of the study was to assess the criterion-related and structural validity of the items which had been developed in Phases 1 and 2.

Method

*Participant*s. A sample of 693 respondents took part in this study (55% = males and 45% = females). Of these 44% were coaches (*M* age = 44.24, *SD* = 11.51) and 56% were athletes (*M* age = 24.86, *SD* = 7.58). Participants were recruited from a wide range of both individual (e.g., athletics, cycling, swimming) and team sports (e.g., football, netball, cricket), as well as a range of competitive levels including recreational (2.7%), University (10.2%), club (31.6%), regional (22.5%), national (17.3%), and international (15.7%). Participants had been involved in their primary sport for a mean of 10.17 years (*SD* = 8.07). The average length of their relationship with their current coach or athlete was 3.32 years (*SD* = 3.42), with the mean number of hours being spent with this person in training each week being 4.91 hours (*SD* = 4.88).

Instrumentation. Two long versions of the *Coach-Athlete Relationship Questionnaires* (CART-Qs) were developed, based on the results of Phases 1 and 2. One

version was designed for the athlete and another for the coach. Both versions contained 128 relationship statements (64 direct items and 64 meta items). These contained 42 items designed to measure closeness (21 direct and 21 meta), 40 for commitment (20 direct and 20 meta), and 46 to measure complementarity (23 direct and 23 meta).

Moreover, three items from the "Overall Performance" sub-scale of the *Elite Athlete Self Description Questionnaire* (Marsh, Hey, Johnson, & Perry, 1997) were employed in both a direct and meta form. Support for the factorial validity and internal consistency of this measure has been shown in previous research (Marsh et al., 1997; Marsh & Perry, 2005). Three items from the "Satisfaction with Performance" sub-scale of the *Athlete Satisfaction Questionnaire, which has been found to show good psychometric properties* (Riemer & Chelladurai, 1998) were also included. All items began with the prefix "During training…" such that participants focused on their sporting relationship with their coach or athlete. Respondents indicated the extent to which they agreed with each statement on a 7 point scale ranging from 1 = *Strongly Disagree* through to 7 = S*trongly Agree*.

Procedures. In order to achieve the relatively large and heterogeneous sample required for this study, a wide range of different strategies were employed to recruit participants. This approach is in line with Duncan, Strycken, Duncan, and Chaumenton (2002) who argued that using a range of recruitment methods is advantageous because it increases the success of the overall project. Firstly, National organizations, such as National Governing Bodies (NGBs) from a wide range of sports were contacted via e-mail and/or telephone to invite them to participate in the present study through providing access to coaches and athletes. Secondly, clubs, groups and societies were approached via their head coach/manager. Finally, participants were recruited on an individual level through attending sporting events, courses and training sessions.

Once potential participants had been identified, they were provided with an information pack, either by e-mail, post or through face-to-face contact. This pack contained a letter which introduced the aim of the overall research and a questionnaire. Participants were reassured that any information which they provided would remain confidential at all times. And were made aware that their participation was completely voluntary and that they were free to withdraw from the study at any time. Where appropriate, participants were also provided with a free post envelope. Ethical clearance for the study was obtained from the University's Ethical Advisory Committee.

Data Analysis. Descriptive statistics such as means, standard deviations, bivariate correlations, and internal consistency scores of the performance and satisfaction variables, as well as for the longer sub-scales of the CART-Q were computed. Structural Equation Modeling (SEM) which has become a widely used statistical tool in the social sciences (Martens, 2005), was employed and Confirmatory Factor Analysis (CFA) using EQS 6.1 for Windows (Bentler & Wu, 2005) was applied to test whether the data from the long versions of the CART-Qs fit the hypothesized 3 Cs conceptualization. The use of CFA when assessing construct validity is advocated within this case as it benefits from adjusting for measurement error.

Results

Table 1 displays the means and standard deviations of the long versions of the CART-Qs (direct and meta) and their correlations with performance and satisfaction. Relatively high mean scores were found for all of the relationship variables indicating that many of the respondents viewed their coach-athlete relationship positively.

insert Table 1 around here

Prior to pooling the coach and athlete data, and the male and female data, the heterogeneity of the associated covariance matrices was assessed following the procedures outlined by Byrne and Stewart (2006). In all cases, the differences between Comparative

Fit Indices scores was less than 0.01 and between χ^2 scores was less than 0.05 which indicates that there are no differences in the latent structures of these sub-groups (Cheung & Rensvold, 2002). The data could therefore be combined to create a single data set.

Structural Validity. Many researchers have advocated that a number of goodness-of-fit indices are employed in order to evaluate model fit (Marsh, 2007; Hu & Bentler 1998). Therefore the present study utilized four different statistics: the Comparative Fit Index (CFI), the Bentler-Bonnet Non-Normed Fit Index (NNFI), the Standardized Root Mean-Square Residual (SRMR) and the Root Mean-Square Error of Approximation (RMSEA). For a model to be seen as having an acceptable fit, it should have a CFI and an NNFI of at least 0.90 (Marsh, 2007; Kline, 2005) and a SRMR of less than 0.10 (Kline, 2005). Browne and Cudeck (1993) suggest that an RMSEA between 0 and 0.05 indicates a close fit, less than 0.08 represents a reasonable fit and greater than 0.08 suggests a poor fitting model. The sample of 693 ensured that, in all of the CFA analyses, the participant/item ratio was greater than 10:1 (as recommended by Everitt, 1975) and that Bentler's (1995) guideline of having a free parameter-item ratio also greater than 10:1 was also met. In general terms, a sample size of over 500 is said to be very good for the purposes of CFA (Comrey & Lee, 1992). This relatively large sample should therefore help minimize the risks of patterns of co-variation being unstable and the sample not representing the target population (DeVellis, 2003).

The two-stage approach advocated by Anderson and Gerbing (1988) was used. In the first stage of the analysis, the item pools for closeness (21 items), commitment (20 items) and complementarity (23 items) were analyzed independently using CFA. The modification process was theory-driven and fundamentally based on the 3+1C conceptualization (Jowett, 2005, 2007a). Thus items were only retained if they loaded onto the factor which they had been developed to measure. The modifications which were grounded in theory were supported by the modification indices provided by EQS 6.1 for

Windows (Bentler & Wu, 2005), namely Lagrange Multiplier and Wald test. Items which did not load on to the factor to which they were conceptually associated were identified and deleted one by one and a CFA was conducted on the resultant items until satisfactory fit indices were achieved for both the direct and meta versions of the long CART-Q. This process resulted in three final scales designed to measure closeness (10 items), commitment (11 items) and complementarity (15 items).

The second stage of the analysis involved assessing whether the 3+1C conceptualization satisfactorily fit the data. Two different models were tested. Model 1 (M1) hypothesized a general higher order factor of relationship quality which accounts for the correlations between the three first order factors of closeness, commitment and complementarity. In Model 2 (M2), a three first order factor model was tested comprising of closeness, commitment and complementarity. These two models were selected as they were found to have the best fit indices in the validation of the original CART-Qs (Jowett, 2007b; Jowett & Ntoumanis, 2004).

The robust maximum likelihood estimation procedure was employed due to a relatively high normalized estimate of Mardia's coefficient (multivariate kurtosis = 90.87) as recommended by Bentler (1995). This approach adjusts the standard errors and the Chi Square statistic under conditions of non-normality in order to protect against the risk of a type 1 error. Theory-driven modifications, supported by the Lagrange Multiplier and Wald tests, were made. Such refinements are a common requirement in the development of instruments (MacCallum, Roznowski, & Necowski, 1992). The model fit was re-evaluated after each item had been removed. Hoffman (1995) supports this process as it has the benefit of maintaining the general structure of the hypothesized factor model with the best indicators. This process resulted in a final version of the questionnaires which contains 29 items (direct and meta): 7 for closeness, 10 for commitment and 12 for complementarity.

(The items of the long versions of the direct and meta-perspective of the CART-Q are available from the authors).

For the direct items, the fit indices achieved the recommended cut-off points (see Table 2). Model 1, which hypothesized a higher order factor did not yield as good indices as Model 2. For Model 2, which purports three first order factors, the SRMR of .07 is below the recommended .10 cut-off point. Furthermore, the RMSEA of .07 indicates a reasonable model fit. Also, both the RCFI and the NNFI achieved adequate fit indices of .9. These statistics were supplemented by the Akaike Information Criteria (AIC) which assesses if fewer estimated parameters could be used to achieve a satisfactory model fit. Although no specific cut-off values are provided for use with this statistic, the model with the lowest AIC value is suggested to provide the more parsimonious fit to the data. Therefore, as the AIC is lower for M2 than for M1, further evidence is provided for the conclusion that M2 achieves a better fit to the data than M1. Thus, the M2 model with three first order factors which are taken to represent closeness, commitment and complementarity has been shown to approximate an adequate fit to the data. The factor loadings and error variances for closeness ranged from .56 - .91 (.31 - .64), for commitment ranged from .74 - .94 (.34 - .70), and for complementarity ranged from .75 - .88 (.47 - .70). All loadings were statistically significant at the 0.05 level. The inter-factor correlations were as follows: .51 for closeness and commitment, .49 for closeness and complementarity, and .19 for commitment and complementarity.

insert Table 2 near here

Similar findings were found for the meta items with the three first order factor model (M2) again fitting the data more accurately than the higher order factor model (M1) and also having a lower AIC value (see Table 2). The factor loadings and error variances for meta-closeness ranged from .76 - .91 (.42 - .65), for meta-commitment ranged from .71 - .90 (.41 - .70), and for meta-complementarity ranged from .73 - .91 (.38 - .70). All

loadings were statistically significant at the 0.05 level. The inter-factor correlations were as follows: .53 for meta-closeness and meta-commitment, .51 for meta-closeness and meta-complementarity, and .22 for meta-commitment and meta-complementarity.

Reliability. The sub-scales of the long versions of the CART-Q were all found to have satisfactory internal consistency as they exceeded the cut off point of .70 (Nunnally & Bernstein, 1994). Specifically, Direct Closeness $\alpha = .92$, Direct Commitment $\alpha = .94$, Direct Complementarity $\alpha = .94$, Meta Closeness $\alpha = .91$, Meta Commitment a = .92 and Meta Complementarity $\alpha = .94$.

Criterion Validity. Criterion validity is "…a measure of how well a variable, or set of variables, predict an outcome" (Pennington, 2003, p. 37). In this study, concurrent validity was tested using the criterion variables of satisfaction and subjective performance. There is theoretical and empirical evidence to link coach leadership (see Riemer, 2007) and coach-athlete relationships (e.g., Jowett & Ntoumanis, 2004) with satisfaction and performance. The two sub-scales used as criterion variables were found to be internally consistent: satisfaction $\alpha = .81$ and performance $\alpha = .91$. Results from the bivariate correlations (see Table 1) indicated that the six sub-scales of the long CART-Qs were all significantly correlated with both satisfaction and performance which provides some evidence for the criterion-related validity (concurrent) of the new instruments.

Two hierarchical multiple regression analyses were also conducted. This enabled an assessment of whether the long versions of the CART-Q were able to predict a significantly higher amount of variance in the criterion variables of satisfaction and performance than the short CART-Q. Firstly, satisfaction was predicted from both the direct and meta perspectives of the short and longer versions of the CART-Q. Specifically, the direct 3Cs from the short CART-Q were entered into the first step and the meta perspectives of the 3Cs on the short CART-Q were entered into the second step. The direct and meta perspectives of the 3Cs as measured by the long CART-Q were entered into the third and

fourth steps. A similar approach was conducted in the second regression equation but with performance being the dependent variable.

The direct perspective of the 3Cs of the short CART-Q (R^2 = .29, F = 82.50, p < .001) and the meta 3Cs of the short CART-Q (R^2 = .29, F = 43.08, p < .01) were both able to predict a significant amount of the variance in satisfaction. The addition of the direct 3Cs of the long CART-Q (R^2 = .36, F = 38.17, p < .01) and the meta 3Cs of the long CART-Q (R^2 = .36, F = 30.03, p < .01) both explained a further significant amount of variance in satisfaction. Thus, the long version of the CART-Qs were able to explain 7.1% more variance in satisfaction over and above that accounted for by the short CART-Qs.

Furthermore, the direct perspective of the 3Cs of the short CART-Q (R^2 = .27, F = 77.72, p < .01) and the meta 3Cs of the short CART-Q (R^2 = .30, F = 54.00, p < .01) were both able to predict a significant amount of the variance in performance. The inclusion of the direct 3Cs of the long CART-Q (R^2 = .39, F = 48.36, p < .01) and the meta 3Cs of the long CART-Q (R^2 = .40, F = 28.31, p < .01) both explained a further significant amount of variance in performance. Thus, the long versions of the CART-Qs were able to explain 9.4% more variance in performance over and above that accounted for by the short CART-Qs. These analyses provide empirical evidence that the expanded CART-Qs can provide more information about the coach-athlete relationship which can facilitate a better prediction of a coach or an athlete's satisfaction or perceived performance.

Study 2

The second study aimed to administer the new measure to an independent sample of coaches and athletes. This afforded a further assessment of the structural validity of the longer version of the CART-Q and provided more evidence regarding the psychometric properties of the newly developed instrument.

Method

*Participant*s. A sample of 251 respondents took part in this study (49% = males and 51% = females). Of these 42% were coaches (*M* age = 37.71, *SD* = 10.05) and 58% were athletes (*M* age = 19.82, *SD* = 3.08). Participants were recruited from a wide range of both individual (e.g., athletics, golf and swimming) and team sports (e.g., football, netball and rowing), as well as a range of competitive levels: recreational (2.0%), University (23.8%), club (25.4%), regional (23.8%), national (16.4%), and international (8.6%). Participants had been involved in their primary sport for a mean of 7.78 years (*SD* = 4.46). The average length of their relationship with their current coach or athlete was 2.57 years (*SD* = 2.37), with the mean number of hours being spent with this person in training each week being 4.57 hours (*SD* = 3.65).

Materials. The administered questionnaire contained the 29 item direct perspective version of the long CART-Q (which was developed in Study 1). All items began with the prefix 'During training…' to emphasize that the questionnaire is focused on the respondent's sporting relationship with their coach/athlete. The long CART-Q is comprised of 7 items which measure closeness (e.g., I care about my coach/athlete), 10 items that measure commitment (e.g., I am committed to maintaining a close partnership with my coach/athlete) and 12 items which measure Complementarity (e.g., I am organized). The third section contained the meta perspective of the long CART-Q. This has very similar items to the direct perspective version with the only difference being that they were re-worded to ask the respondent to think about how their coach/athlete feels, thinks and behaves. The meta perspective of the long CART-Q therefore measures meta closeness (e.g., my coach/athlete cares about me), meta commitment (e.g., my coach/athlete is committed to maintaining a close partnership with me) and meta Complementarity (e.g., My coach/athlete is organized). Respondents indicated their agreement with the items on a 7 point scale from 1 'strongly disagree' to 7 'strongly agree'.

Procedure. The same recruitment methods as those used in Study 1 were employed.

Data Analysis. Confirmatory Factor Analyses were conducted using EQS 6.1 for Windows (Bentler & Wu, 2005) to assess the structural validity of the long CART-Qs. Cronbach's Alpha was also used to assess internal consistency.

Results

As in Study 1, the robust maximum likelihood estimation procedure was employed due to a relatively high normalized estimate of Mardia's coefficient (multivariate kurtosis = 36.37) as recommended by Bentler (1995). CFA was employed to assess the extent to which M1 (the first order model taken to represent closeness, commitment and complementarity) fit the data. The data for both versions yielded good model fit indices. The RCFI (direct = .95 and meta = .97), the NNFI (direct = .94 and meta = .96) and the RMSEA (direct = .07 and meta = .07) all met the recommended criteria for an acceptable model fit (Brown & Cudeck, 1993; Kline, 2005). In order to further examine the structural validity of the scale responses the fit of alternative models to the data were also assessed. These included M2 which was a hierarchical model containing the 3Cs and a higher order factor taken to represent the overall quality of the relationship, M3 which was a unidimensional model and M4 which was a 3 factor uncorrelated model (containing the factors of closeness, commitment and complementarity). The results are displayed in Table 3 and these demonstrate that the 3 first order correlated factor model has a superior fit to all of these alternative models which provides further support for the dimensionality and structural validity of the new measure.

Due to the non-normality of the data, the Satorra-Bentler scale chi-square statistics were compared. In each case M1 was found to be superior. When comparing M1 and M2 xdiff = 56.28 ($p < .01$). When comparing M1 with M3 xdiff = 87.86 ($p < .01$) and when comparing M1 with M4 xdiff = 44.73 ($p < .01$). Thus, in all cases there was a significant loss of fit when moving from the 3 correlated first order factor model to any of the alternative models, which lends further support to the structural validity of the measure.

Insert Table 3 near here

Discriminant validity was also tested through comparing the fit of the correlated 3 first order factor model with alternative competing approaches in which all possible pairs of factors are hypothesized to load on to a single factor. The results of these analyses are shown in Table 3. In M5 the two factors were comprised of closeness and another factor in which the commitment and complementarity items were hypothesized to load on to the same factor. In M6 the two factors were comprised of commitment and another factor in which the closeness and complementarity items represented a single factor. In M7 the two factors were comprised of complementarity and another factor in which the commitment and closeness items were hypothesized to load on to the same factor.

As in the above analyses, due to the non-normality of the data, the Satorra-Bentler scale chi-square statistics were compared. In each case M1 was found to be superior. When comparing M1 and M5 xdiff = 26.70 ($p < .01$). When comparing M1 with M6 xdiff = 47.47 ($p < .01$) and when comparing M1 with M7 xdiff = 52.91 ($p < .01$). Thus, in all cases there was a significant loss of fit when moving from the 3 correlated first order factor model to any of the alternative models, which lends further support to the structural validity of the new measure.

Further analyses using Cronbach's alpha provided support for the reliability of the long CART-Qs sub-scales; direct closeness a = .85, direct commitment a = .88, direct Complementarity a = .88, meta closeness a = .87, meta commitment a = .88 and meta Complementarity a = .89.

General Discussion

The current research focused on gaining initial evidence for the criterion validity and the structural validity of the scale's responses on an expanded measure of the quality of the coach-athlete relationship (CART-Q; Jowett, in press, 2007b; Jowett & Ntoumanis, 2004). This means that the CART-Q is now available in both a short and long version.

Phases 1 and 2 of Study 1 involved the generation of items and the assessment of their content validity by expert panels. This ensured that the items contained within the longer CART-Q are viewed as being relevant and representative of the construct of interest.

The structural validity of the longer versions of the CART-Q was assessed in Phase 3. The 3 first order 3 C model (i.e., closeness, commitment and complementarity), adequately fit the data and provided a better fit than the higher order model. This supports the findings of Jowett's (in press; 2007b) research which validated the meta version of the CART-Q. Study 2 lends further support to the structural validity and reliability of the newly developed measure using an independent sample. The 3 correlated first order factor model taken to represent closeness, commitment and complementarity was found to fit the data significantly better than a number of alternative models.

Associations between the expanded measure and performance and satisfaction supported previous research findings (e.g., Jowett & Cockerill, 2003; Jowett & Ntoumanis, 2004) and also provide evidence for the criterion-related validity of the longer CART-Qs. Therefore, initial evidence for the three important forms of validity recommended by the American Psychological Association (1999), content, structural and criterion-related, has been found. All of the sub-scales of the longer versions of the CART-Qs were also found to be internally consistent. Research is needed to assess the stability of the measure over time.

Evidence supporting the expansion of the CART-Q was provided by the hierarchical multiple regression analyses. They indicated that the longer instruments can account for 7.1% of the variance in satisfaction and 9.4% of the variance in perceived performance over and above that explained by the short CART-Qs. This research has an important contribution to make in theoretical, applied, and research terms. By expanding the questionnaire, its scope to include relevant interpersonal aspects has been broadened. For example, the closeness sub-scale now includes items regarding whether a coach and athlete value and care about one another. The commitment sub-scale has been expanded to

include, for example, items regarding whether the coach/athlete can imagine ending the relationship in the next year, how much their sporting success is influenced by the relationship member, and how much has been invested in the relationship. As for complementarity, items regarding being receptive and paying attention have been added along with questions about being organized and clear about what is expected of each relationship member during training sessions.

This new measure could provide information upon which coach education programs can be developed and to ensure that coaches, athletes, sport psychologists, and researchers have a greater knowledge-base upon which to develop their understanding of the nature and role of the coach-athlete relationship. This knowledge should, in turn, help all of these interested parties to take measures to optimize the effectiveness of coach-athlete relationships and hence the performance and satisfaction of the coach and athlete.

With reference to applied sport psychology, the expanded CART-Q represents a tool which will enable sport psychology consultants to obtain a more detailed assessment of the quality of a coach-athlete relationship. The results of such an assessment could provide them with information upon which they can formulate decisions regarding potential areas of strength or weakness in a relationship. Such an assessment could assist toward taking appropriate steps to ensure that the relationship continues at an optimal level into the future. This helps to narrow the gap highlighted by Coppel (1995) in terms of scientific knowledge and its usefulness in applied terms. It also helps to address the need identified by Wylleman (2000) to develop, validate, and delineate measures regarding the coach-athlete relationship.

In research terms, the present study helps to diversify theory and research from other social scientific domains into the field of sport psychology, as recommended by Poczwardowski, Barott, and Jowett (2006). The expanded CART-Q provides researchers with a comprehensive assessment of the quality of a coach-athlete relationship and will

represent a useful resource for researchers intending to investigate the quality of a coach-athlete relationship but for whom the original CART-Q (see e.g., Jowett, 2007b; Jowett & Ntoumanis, 2004) is too brief. Furthermore, the three sub-scales can be used independently and researchers can select a sub-scale/s based on their specific interests.

A number of lines of investigation open up. Firstly, further validation of the longer versions of the CART-Q is merited. The factorial structure of the long versions of the CART-Q need to be assessed with different samples (e.g., team versus individual sports, recreational versus elite performers). There is also great scope for the expanded measure, whether the three sub-scales are used independently or together, to investigate the nature and role of the coach-athlete relationship. Such research could assess the associations of the coach-athlete relationship with individual factors (i.e., the gender and age of the coach and athlete, motivation, passion), relationship factors (e.g., atypical relationships, same-sex and different sex dyads, relationship styles), and environmental factors (e.g., team versus individual sports, different types of sport, and different levels of performance). The long versions of the CART-Q could be used in longitudinal and intervention research.

Research and interest in the coach-athlete relationship is likely to continue to increase into the future. With many sport-related positive outcomes being related to quality relationships (Petitpas, 2002), and with award-winning coaches placing a tremendous emphasis on the role played by quality relationships in terms of both success and personal development (Gould et al., 2007), it is highly probable that the demand for information, conceptualizations and measurements regarding the coach-athlete relationship will also increase. The expanded CART-Q can help to meet this demand through facilitating education, research and applied sport psychologists in the drive towards optimizing sport performance and participant satisfaction.

References

American Psychological Association (1999). *Standards in Educational and Psychological Testing (4th Ed.)*. Washington, DC: American Psychological Association.

Anastasi, A. (1988). *Psychological testing* (6th Ed.). New York: Macmillan.

Anderson, J. C., & Gerbing, D. W. (1988). Structural equation modeling in practice: A review and recommended two-step approach. *Psychological Bulletin, 3,* 411–423.

Bentler, P.M. (1995). *EQS structural equations program manual*. Encino, CA: Multivariate Software.

Bentler, P. M. & Wu, E. J. C. (2005). *EQS 6.1 for Windows*. Encino, CA: Multivariate Software (Context Link).

Browne, M. W., & Cudeck, R. (1993). Alternative ways of assessing model fit. *Sociological Methods and Research, 21*, 230-258.

Byrne, B. M., & Stewart, S. M. (2006). The MACS approach to testing for multi-group invariance of a second order structure: A walk through the process. *Structural Equation Modelling, 13*, 287-321.

Cheung, G. W., & Rensvold, R. B. (2002). Evaluating goodness-of-fit indices for testing measurement invariance. *Structural Equation Modelling*, 9, 233-255.

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

Coppel, D.B. (1995). Relationship issues in sport: A marital therapy model. In S.M. Murphy (Ed.), *Sport psychology interventions* (pp. 193-204). Champaign, IL: Human Kinetics.

DeVellis, R. F. (2003). *Scale development: Theory and applications (2nd Ed.).* CA: Sage.

Duda, J. L. (2001). Achievement Goal Research in Sport: Pushing the boundaries and clarifying some misunderstandings. In G. C. Langer (Ed.) *Advances in motivation in sport and exercise* (pp. 129-182). Champagne, IL: Human Kinetics.

Duncan, S.C., Strycker, L.A., Duncan, T.E., & Chaumenton, N.R. (2002). Telephone recruitment of a random stratified youth sample for a physical activity study. *Journal of Sport & Exercise Psychology, 24,* 347-358.

Everitt, G. S. (1975). Multivariate Analysis: The need for data and other problems. *British Journal of Psychiatry, 126,* 237-240.

Gill, D. L., Dzewaldowski, D. A., & Deeder, T. E. (1988). The relationship of competitiveness and achievement orientation to participation in sport and non-sport activities. *Journal of Sport and Exercise Psychology*, *10,* 139-150

Gould, D., Collins, K. B., Louer, L. A., & Chung, Y. C. (2007). Coaching life skills through football: A study of award-winning high-school coaches. *Journal of Applied Sport Psychology, 19,* 16-37.

Hoffmann, R. (1995). Establishing factor validity using variable reduction in confirmatory factor analysis. *Educational and Psychological Measurement, 55,* 572–582.

Hoyt, W. T., Warbasse, R. E., & Chu, E. Y. (2006). Construct validation in counseling psychology research. *The Counseling* Psychologist, *34*, 769-792.

Hu, L., & Bentler, P. M. (1998). Fit indices in covariance structure modeling: Sensitivity to under-parameterised model mis-specification. *Psychological Methods, 3,* 424-453.

Jowett, S. (2003). When the honeymoon is over: A case study of a coach - athlete relationship in crisis. *The Sport Psychologist, 17,* 444-460.

Jowett, S. (2005). On repairing and enhancing the coach-athlete relationship. In S. Jowett & M. Jones (Ed.) *Psychology of sport coaching*. Leicester: British Psychological Society.

Jowett, S. (2007a). Closeness, commitment, complementarity and co-orientation in the coach-athlete relationship: An interdependence analysis. In S. Jowett & D. Lavallee (Eds.) *Social psychology in sport* (pp. 15-28). Champagne, IL: Human Kinetics.

Jowett, S. (2007b). Factor Structure and Criterion Validity of the Meta-Perspective Version of the Coach-Athlete Relationship Questionnaire (CART-Q). *Manuscript under review*.

Jowett, S. (in press). Validating coach-athlete relationship measures with the nomological network. *Measurement in Physical Education and Exercise Science.*

Jowett, S., & Chaundy, V. (2004). An investigation into the impact of coach leadership and coach-athlete relationship on group cohesion. *Group Dynamics: Theory, Research and Practice, 8,* 302–311.

Jowett, S., & Cockerill, I. M. (2003). Olympic Medalists' perspective of the athlete - coach relationship. *Psychology of Sport and Exercise, 4,* 313-331.

Jowett, S., & Clark-Carter, D. (2006). Perceptions of empathic accuracy and assumed similarity in the coach-athlete relationship. *British Journal of Social Psychology, 45,* 617-637.

Jowett, S., & Meek, G. (2000). Coach – athlete relationships in married couples: An exploratory content analysis. *The Sport Psychologist, 14*, 157-175.

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

Jowett, S., & Wylleman, P. (2006). Interpersonal relationships in sport and exercise: Crossing the chasm. *Psychology of Sport and Exercise, 7,* 119-123.

Kelley, H. H., Berscheid, E., Christensen, A., Harvey, H. H., Huston, T. L., Levinger, G., et al. (Eds.) (1983). *Close relationships.* New York: Freeman.

Kiesler, D.J. (1997). *Contemporary interpersonal theory research and personality, psychopathology, and psychotherapy*. New York: Wiley.

Kline, R. B. (2005). *Principles and practice of structural equation modeling*, London: Guilford Press.

MacCallum, R. C., Roznowski, M., & Necowitz, L. B. (1992). Model modifications in covariance structure analysis: The problem of capitalization on chance. *Psychological Bulletin, 111,* 490-504.

Marsh, H. W. (2007). Application of confirmatory factor analysis and structural equation modeling in sport/exercise psychology. In G. Tenenbaum & R. C. Eklund (Eds.), *Handbook of sport psychology (3rd Ed.).* New York: Wiley.

Marsh, H. W., Hey, J., Johnson, S., & Perry, C. (1997). Elite Athlete Self-Description Questionnaire: Hierarchical confirmatory factor analysis of responses by two distinct groups of elite athletes. *International Journal of Sport Psychology, 28,* 237-258.

Marsh, H. W., & Perry, C. (2005). Self-concept contributes to winning gold medals: Causal ordering of self-concept and elite swimming performance. *Journal of Sport and Exercise Psychology, 27,* 71-91.

Martens, M. P. (2005). The use of structural equation modeling in counseling psychology research. *The Counseling Psychologist, 33*, 269-298.

Messick, S. (1989). Validity. In R. L. Linn (Ed.), *Educational measurement* (3rd Ed., pp. 13-103). New York: American Council on Education.

Nunnaly, J. C., & Bernstein, I. H. (1994). *Psychometric theory (3rd Ed.).* New York: McGraw-Hill.

Pennington, D. C. (2003). *Essential personality*. London: Arnold.

Petitpas, A. (2002). Counseling interventions in applied sport psychology. In L. Van Raalte & B. W. Brewer (Eds.), *Exploring sport and exercise psychology* (pp. 253-268). Washington, DC: American Psychological Association.

Poczwardowski, A., Barott, J. E., & Jowett, S. (2006). Diversifying approaches to research on athlete-coach relationships. *Psychology of Sport and Exercise, 7*, 125-142.

Poczwardowski, A., Sherman, C., & Henschen, K. (1998). A sport psychology service delivery heuristic: Building on theory and practice. *The Sport Psychologist, 12*, 192-208.

Quintana, S. M., & Minami, T. (2006). Guidelines for meta-analyses of counseling psychology research. *The Counseling Psychologist, 34*, 839-876.

Riemer. H. A. (2007). Multidimensional model of coach leadership. In S. Jowett & D. Lavallee (Eds.) *Social psychology in* sport (pp. 57-74). Champagne, IL: Human Kinetics.

Riemer, H. A. & Chelladurai, P. (1998). Development of the Athlete Satisfaction Questionnaire (ASQ). *The Journal of Sport and Exercise Psychology, 20*, 127-156.

Rubin, Z. (1970). Measurement of romantic love. *Journal of Personality and Social Psychology, 16*, 265-273.

Rusbult, C.E., Martz, J. M., & Agnew, C. R. (1998). The Investment Model Scale: Measuring commitment level, satisfaction level, quality of alternatives, and investment size. *Personal Relationships, 5*, 357-391.

Wylleman, P. (2000). Interpersonal relationships in sport: Uncharted territory in sport psychology research. *International Journal of Sport Psychology, 31*, 555-572.

Table 1

Means and Standard Deviations of the extended 3Cs and correlations with performance and satisfaction

Variables/Subscales	*Mean*	*SD*	Direct Satisfaction	Direct Performance
Direct Closeness	5.97	0.93	.46*	.35*
Direct Commitment	5.13	1.45	.54*	.74*
Direct Complementarity	6.11	0.90	.25*	.23*
Meta Closeness	5.69	0.97	.46*	.34*
Meta Commitment	4.99	1.43	.48*	.70*
Meta Complementarity	6.01	0.94	.24*	.22*

* $p < .05$

Table 2

Fit indices for the higher order and first order models for the direct and meta items

Model	RCFI	NNFI	SRMR	RMSEA	90% sig. CI	AIC
M1 (direct)	.87	.86	0.22	0.10	.095-.103	932.36
M2 (direct)	.90	.90	.07	0.07*	.066-.074	759.95
M1 (meta)	.84	.82	.23	0.11	.106-.114	867.26
M2 (meta)	.90	.90	.08	0.07*	.066-.074	736.41

Note. RCFI = Robust Comparative Fit Index, NNFI – Non-Normed Fit Index, SRMR – Standardized Root Mean-Square Residual, RMSEA = Root Mean-Square Error of Approximation, 90% CI = 90% Confidence Interval of RMSEA and AIC = Akaike Information Criterion, * = not significant

Table 3

Fit indices for alternative models for the direct and meta items

Model	RCFI	NNFI	SRMR	RMSEA
M1 (direct)	.95	.97	0.07	0.07*
M1 (meta)	.94	.96	.07	0.07*
M2 (direct)	.85	.84	.23	0.12
M2 (meta)	.82	.82	.25	0.13
M3 (direct)	.70	.77	.28	0.20
M3 (meta)	.77	.76	.24	0.18
M4 (direct)	.81	.81	.16	0.19
M4 (meta)	.83	.82	.21	0.24
M5 (direct)	.88	.87	.11	0.10
M5 (meta)	.87	.85	.13	0.11
M6 (direct)	.85	.85	.16	0.15
M6 (meta)	.86	.87	.12	0.14
M7 (direct)	.83	.83	.13	0.13
M7 (meta)	.84	.83	.14	0.12

Note. RCFI = Robust Comparative Fit Index, NNFI – Non-Normed Fit Index, SRMR – Standardized Root Mean-Square Residual, RMSEA = Root Mean-Square Error of Approximation, * = not significant