

Running head: SHARED BELIEFS

Do Individual Perceptions of Group Cohesion Reflect Shared Beliefs?

An Empirical Analysis

Albert V. Carron¹, Lawrence R. Brawley², and Mark A. Eys¹

and in alphabetical order

Steven Bray³, Michelle Colman¹, Kim Dorsch⁴, Paul Estabrooks⁵, Craig R. Hall¹, James Hardy¹, Heather Hausenblas⁶, Ralph Madision⁸, David Paskevich⁷, Harry Prapavessis⁸,
Kevin S. Spink⁹, & Peter C. Terry¹⁰

¹University of Western Ontario, Canada; ²University of Waterloo, Canada; ³University of Lethbridge, Canada; ⁴University of Regina, Canada; ⁵Kansas State University, ⁶University of Florida, ⁷University of Calgary, Canada; ⁸University of Auckland, New Zealand; ⁹University of Saskatchewan, Canada; ¹⁰University of Southern Queensland, Australia

Submitted: September 3, 2002

Address Mail Correspondence to:

Albert V. Carron

School of Kinesiology

University of Western Ontario

London, Ontario, N6A 3K7, CANADA

Address Email Correspondence to: bcarron@uwo.ca

Running head: SHARED BELIEFS

Do Individual Perceptions of Group Cohesion Reflect Shared Beliefs?

An Empirical Analysis

Abstract

The primary purpose of the study was to use the index of agreement to examine the degree to which individual perceptions of cohesiveness reflect shared beliefs for group cohesion in sport teams. The secondary purposes were to examine how the type of cohesion, the task interactive nature of the group, and the absolute level of cohesion are related to the index of agreement. Teams ($n = 192$ containing 2107 athletes) were tested on the Group Environment Questionnaire. The average index of agreement values were greater for the Group integration (GI) manifestations of cohesiveness (GI-task, $r_{wg(j)} = .721$; GI-social, $r_{wg(j)} = .694$) than for the Individual attractions to the group (ATG) manifestations (ATG-task, $r_{wg(j)} = .621$; ATG-social, $r_{wg(j)} = .563$). No differences were found in index of agreement values for teams differentiated on the basis of task interdependence (i.e., interactive sport teams versus coactive/independent sport teams). Finally, a consistent pattern of relationships was observed between the absolute level of cohesiveness and the index of agreement; consensus increased as the amount of cohesiveness present in the team increased. The results were discussed in terms of their implication for the aggregation of individual perceptions of cohesion to represent the group construct.

KEYWORDS: Task type; Consensus; Sport teams; Group Environment Questionnaire

Do Individual Perceptions of Group Cohesion Reflect Shared Beliefs?

An Empirical Analysis

In his contemporary review of group cohesion, Dion (2000) provided a selective historical tour of the rich history of multidisciplinary investigations associated with the construct of group cohesion. He noted that social psychology, as one of the investigative disciplines, brought its distinctive focus upon conceptualization, measurement and validity issues to the examination of the cohesion phenomenon. As he correctly emphasized, these issues continue to be important if the research on cohesion is to advance. The objective of the present article is to focus on one such issue—the notion that cohesion, when conceptualized and measured through the perceptions of individuals within the group, represents the shared beliefs of the membership. As Dion (2000) has aptly noted, cohesion is clearly important in sport and exercise groups. In the present study, naturally-formed groups are used to quantitatively illustrate the degree of sharing that occurs in beliefs about cohesion. In order to appreciate the importance of this demonstration, some background about shared beliefs in groups, its link to cohesion, and its statistical representation is in order.

Shared Beliefs and Group Cohesion

The notion that members of a group share beliefs about matters of importance has been linked to the investigation of cohesion for some time. For example, in their classic study of housing units, Festinger, Schacter, and Back (1950) reasoned that if the group used its cohesion to influence its members to think and act alike, then cohesion should be related to group homogeneity of attitudes and behavior. Another example is found in cohesion studies of Janis's (1982) "groupthink" concept that stresses the single mindedness of thinking (i.e., sharing) that is characteristic of some groups (e.g., not willing to consider perspectives other than the group's). Groupthink has been shown to be

both encouraged and countered by different types of cohesion (Bernthal & Insko, 1993). Both of these examples reflect the idea of individual members of small groups holding common (shared) perceptions. Not surprisingly, investigators have assumed that members' responses would be similar within more cohesive groups and that this commonality can be measured.

The social psychological literature is replete with different ways to measure cohesion (cf., Dion, 2000). One approach has been to measure the perceptions of individual group members about the degree to which various dimensions of cohesion are present in the group. Then, depending upon the nature of the research question, the individual responses have been aggregated and the resultant measures then assumed to reflect the cohesion of the entire group. Dion (2000) described two programs of research that exemplify this approach. For example, Bollen and Hoyle (1990) developed a two-dimensional conceptualization of cohesion that concerns belongingness and morale. In commenting on their approach, Dion (2000) noted that at the individual's level, perceived cohesion represents the group's role for its members. Furthermore, combining the individual perceptions represents a group level view of cohesion. Perceived cohesion, from the Bollen and Hoyle viewpoint, applies to both small and large groups.

Dion (2000) also described another approach to measuring perceived cohesion in sport groups. Four aspects of cohesion have been proposed by Carron, Widmeyer, and Brawley (1985) within their cohesion inventory, the Group Environment Questionnaire (GEQ). Considerable sport-related evidence for the construct validity of the GEQ has accumulated since its development (Carron, Brawley, & Widmeyer, 1998). A fundamental assumption of the Carron et al. approach is that "perceptions about the group held by a group member are a reasonable estimate of various aspects of unity characteristic of the group" (Carron et al., 1998, p.217). Another basic assumption is that

these beliefs are shared. Group (team) members are given 18 statements pertaining to 4 manifestations of task or social cohesiveness and asked to indicate their level of agreement on a 9-point scale. Typically, group member responses to the GEQ have been used either independently (e.g., to examine relationships between individual perceptions of group cohesion and individual adherence behavior; Carron, Widmeyer, & Brawley, 1988) or have been aggregated to provide a group value (e.g., to examine relationships between team cohesion and team success; Carron, Bray, & Eys, 2002). Both the Bollen and Hoyle (1990) and the Carron, Widmeyer and Brawley (1985) measures of cohesion have the common assumption that beliefs about cohesion are shared by group members.

Sharing of Information in Groups

Whether individual perceptions about cohesion represent shared beliefs could be questioned. The idea of individuals sharing cognition or sharing beliefs about other aspects of group life frequently refers to both divisions of knowledge shared by members of a group and having some ideas in common. In the former case, group members might have a common goal but divide the cognitive work into different areas of responsibility so that workload is shared in the group's striving toward its common goal. In the latter case, there is a common sense notion that members of the same group will share thoughts and beliefs because of common knowledge and experience. However, as Cole (1991) noted, no one knows everything about their own group or their own culture.

A reasonable question pertains to how much sharing of thoughts and beliefs occur within cohesive groups. The answer to this question may be that "it depends". According to a review by Hinsz, Tindale and Vollrath (1997), task characteristics, group operating procedures, individual member background characteristics, member roles, and group norms all influence the information processing within a group, and as a consequence, the degree of sharing of information within a group. As noted by Levine,

Resnick, and Higgins (1993), unless group members have a shared frame of reference for some information-processing objective, each group member may process information differently. These factors could influence either an increased sharing of information or a heightened variability of individual member information processing.

In support of the idea of shared beliefs is evidence from the comparison of group and individual-information processing (Hinsz et al, 1997). Groups appear to use their information processing strategies more consistently and reliably than individuals. As well, groups are more reliable in the cues they use to make their judgements—whether those judgments are correct or incorrect.

Given the evidence on the various factors that moderate the degree of sharing, reasonable issues to investigate are a) whether groups always reflect the common sharing of beliefs frequently associated with group cohesion and b) whether this sharing is a matter of degree. Donnelly, Carron, and Chelladurai (1978) have noted that by definition, all groups have some degree of group cohesion otherwise they would not exist as groups. Thus, shared beliefs may not necessarily be reflected solely by completely common responses to questions about a group's cohesion.

The idea that there may be a range in the degree of sharing in beliefs about cohesion should also be considered in what we expect from the data gathered with different measures of cohesion. As Dion (2000) has pointed out, there are several theoretically-driven measures of cohesion that attempt to capture its multiple dimensions. He also noted that cohesiveness could mean different things for different groups depending upon the goals and tasks of the group. Therefore, high expectations for commonality of responses to measures reflecting multiple dimensions of cohesion might not be realistic. Finally, as Hinsz et al (1997) illustrated, even the response mode, measure or type of scale affect group responses. For example, expecting within-group

consensus on dichotomous response scales about a single construct (e.g., yes/no; guilty/not guilty) may be probable. Conversely, group consensus on multiple dimensions of a construct assessed with a response scale that has interval or continuous response properties (e.g., 1 = strongly disagree to 9 = strongly agree) may be less probable.

The degree of sharing becomes critical to important decisions about whether to aggregate measures of individual member perceptions of cohesion to the group level (i.e., for research questions that concern the group as the unit of analysis). Many authors have argued that a demonstration of agreement among group members regarding their responses or judgements about their group is essential before stating that a concept is a “group level” phenomenon (e.g., James, Demaree, & Wolf, 1992; Moritz & Watson, 1998). Indeed, Guzzo, Yost, Campbell, and Shea (1993) have noted that low consensus (i.e., minimal shared beliefs) on a topic may be an indication that a group is not of consequence to its members or that it may reflect that the unity of the group is in doubt. Also, Moritz and Watson (1998) have suggested that groups be excluded from group analyses when they have “insufficient intermember agreement” (p. 296). Thus, the demonstration of agreement prior to aggregating individual perceptions about group cohesion is important. However, what is not clear is what level of agreement is sufficient to conclude that shared beliefs are present.

The Index of Agreement

Over the past 20 years, a number of statistical protocols have evolved that permit researchers to empirically test the degree of sharing reflected by the consensus of group members (see Moritz & Watson, 1998, for an overview). One of these statistical protocols is the *index of agreement* ($r_{wg(j)}$, James, 1982; James, Demaree, & Wolf, 1984; Kozlowski & Hatrup, 1992; see also Schmidt & Hunter, 1989). Essentially, the index of agreement provides a statistical measure of the degree to which a set of individuals (e.g.,

a team) is in consensus about a target stimulus (e.g., an item about the team's level of cohesiveness). Mathematically, $r_{wg(j)}$ is defined as a "proportion of systematic variance in a set of judgements in relation to the total variance in the judgments" (James et al., 1984, p. 86).

Kozlowski and Hattrup (1992) pointed out that some researchers have tended to (inappropriately) use the terms interater reliability and interrater agreement interchangeably. They cautioned that this confuses consistency (reliability) with consensus (agreement). Reliability is correlational in nature and reflects the proportional consistency among raters. Thus, for example, one member of a triad might respond with a 9 (on a 9-point scale) for all 4 cohesion items in a scale, another member might respond with a 5 for all 4 items, while the third might respond with a 1. Although reliability (i.e., internal consistency) would be perfect--because raters were consistent across items and the differences among raters were proportionally consistent--agreement would be nonexistent--because no consensus was present among the three raters. The converse also could be true. That is, if all three members of the triad responded to all 4 items with a 7, internal consistency would be zero but consensus would be perfect.

If the index of agreement can be used to assess shared beliefs, the question that then arises pertains to the magnitude of $r_{wg(j)}$ necessary to support the assumption of within-group consensus. As is the case with every other type of statistical test, there is no simple answer. In their overview of statistical procedures used to test for the presence of "groupness" in individual data, Moritz and Watson (1998) offered this suggestion:

if one adopts the James et al. (1984) approach, in which group members are viewed as multiple judges rating the same stimulus, then guidelines from classical test theory can be applied to inform the judgment of whether an observed level of interrater agreement is sufficient. Nunnally

(1978) stated that the sufficiency of a measure's reliability depends on the use to which the measure is put and that, generally, reliability of .50 and .80 can be quite sufficient for research purposes but coefficients in excess of .90 may be required for certain applied purposes. (p. 291)

Moritz and Watson are to be commended for advancing a prescription but one limitation in its application, however, is that the index of agreement ($r_{wg(j)}$) is *not* a reliability coefficient. Thus, any prescriptions pertaining to an acceptable magnitude that might be applicable to statistics such as Cronbach alpha (or intraclass correlations, Spearman-Brown correlation coefficients, test-retest reliability coefficients, etc.) are not necessarily pertinent for the index of agreement.

The primary purpose of the study was to use the index of agreement ($r_{wg(j)}$) to examine the degree to which individual perceptions of cohesiveness reflect shared beliefs for group cohesion in sport teams. A secondary purpose was to examine the influence of various group factors on the index of agreement; the specific factors are discussed below.

In order to reliably establish the range and magnitude of shared beliefs in any construct, four conditions must be present. Two of these pertain to the conditions necessary to calculate an index of agreement (James et al., 1984). First, the inventory must "have acceptable psychometric properties (e.g., construct validity, internal consistency in the case of multiple items) in prior research ... [and] the alternatives on an item's measurement scale [must be] approximately equally spaced" (James et al., 1984, p. 85). The Group Environment Questionnaire (GEQ), the inventory used to test team cohesion in the present study, satisfies these two conditions (see Carron, Brawley, & Widmeyer, 1998).

Two other conditions are related to the nature of the sample. That is, in order to have confidence in the results, the sample should be large and heterogeneous in order to

eliminate questions associated with both power (e.g., are the results simply a reflection of the small number of teams tested?) and generalizability (e.g., are the results simply a reflection of the type of sport examined?). The sample used in the present study--192 teams, comprised of 2107 athletes, from a broad cross section of interactive, independent, and coactive sports—satisfies these two conditions.

Insofar as the secondary purpose was concerned, the influence of three factors on the index of agreement was examined. These included the type of cohesion measure, the nature of the group, and the nature of the group's absolute cohesiveness.

Type of Cohesion

The Group Environment Questionnaire (GEQ, Brawley et al., 1987; Carron et al., 1985; Widmeyer et al., 1985), the operational definition of cohesion used, assesses two predominant types of cognitions. One of these, *group integration*, represents the individual's perceptions about the closeness, similarity, and bonding within the group as a whole. Thus, the items pertain to issues associated with “we” and “us”. The other, *individual attractions to the group*, represents the individual's perceptions about personal motivations acting to retain him or her in the group. Thus, the items pertain to issues associated with “I” and “me”. There are also two fundamental orientations in a group member's perceptions: *task* and *social*. Thus, the GEQ assesses four manifestations of cohesion in sport teams: Group integration-task (GI-T), Group integration-social (GI-S), Individual attractions to the group-task (ATG-T), and Individual attractions to the group-social (ATG-S). After reviewing a number of different conceptualisations of cohesion, Dion (2000) concluded that task and social dimensions were central to a multidimensional view of cohesion and that this view likely reflected a consensus among researchers.

Sport teams are by their very nature task oriented. Most teams practice and prepare for competitions from one to two hours daily and compete at least once a week. Thus, although social interactions are prevalent and important in all groups including sport teams, the major focus and majority of interactions in sport are task-related. Also, as indicated above, the two group integration constructs (GI-T and GI-S) require a judgment about a general state of the group while the individual attractions constructs (ATG-T and ATG-S) require a judgment that is more personal and idiosyncratic. Also, the task interactions are more of a predominant group concern than social interactions. Thus, it was hypothesized that the construct showing the strongest shared beliefs would be GI-T and the construct showing the weakest would be ATG-S.

No hypotheses were advanced about the relative ranking of the other two cohesion constructs—ATG-T and GI-S. This was primarily because the aspects of cohesion can vary in how strongly they are developed depending, for example, upon a number of factors such as the level of team development, the tenure of membership on the team, the type of member task interaction demanded by the sport (i.e., coactive or interactive) and the size of the team (Carron & Brawley, 2000). In a sample as large and diverse as that utilized in the present study, it seemed that a priori hypotheses about the aspects of cohesion where the greatest and least sharing would be present were tenable.

Nature of the Group

Sports differ in the degree to which task interdependence is permitted or required for team success. At one end of a continuum lie independent sport teams. Thus, for example, while high schools, universities and countries routinely enter track and field *teams* in competitions, the rules of the sport mandate that every event must be carried out by a sole individual except for relay races where the nature and amount of interdependence is clearly stipulated. Also, a team score is the additive product of the

success of each individual athlete. Coactive sport teams fall farther along the continuum of interactions. Thus, for example, in rowing teams, members perform coordinative actions but the nature of the task requires a simultaneous synchronized response. Conversely, in sports such as basketball, team members are mutually dependent on one another, task interactions of a variable changing nature, are inherent, and greater cooperation and coordination typically result in better performance.

Shared beliefs should be stronger in situations where more complex task interdependence is fundamental to the nature of the group's success. Thus, it was hypothesized that the stronger shared beliefs would be evident in interactive team sports (e.g., volleyball, basketball) than in coactive/independent sports (e.g., swimming, wrestling).

Absolute Level of Cohesion in the Group

The index of agreement reflects the relative degree of consensus present in member's perceptions of the group. High consensus could result if almost all members perceived that the group had high cohesion, or moderate cohesion, or low cohesion. Cohesion at either of the two extremes is likely a product of unusual distinctive situations—a catastrophe, consistent success, consistent failure, objections to the group's leadership, and so on. Cohesion of a moderate level is likely a product of more ambiguous situations—partial success, inconsistent leadership, and so on. Research in human perception has shown that stimuli or situations that are more distinctive are more noticeable (Hinsz et al, 1997; Kahneman, 1973; Treisman & Gelade, 1980; Treisman, 1988). By extension, it could be expected that consensus should be greater in situations where groups are relatively high or low in cohesiveness than in situations that are more ambiguous. Thus, it was hypothesized that the index of agreement would be greater when team cohesion was highest and lowest than when it was at a moderate/intermediate level.

Methods

Participants

GEQ cohesion data files from multiple investigations were combined to produce the sample. These data files originally were collected for research projects leading to either refereed publications or dissertations/theses. Generally, the studies were carried out under the supervision of or in collaboration with one of the two principal authors¹. Data were obtained from active intact teams (i.e., versus individuals representing teams or recollecting team cohesion). In total, cohesion data considered as collected and scored in reliable fashion as recommended by Carron et al, (1985, 1998, 2002) were obtained for 2107 athletes from 192 teams. Any data not meeting these criteria were not included in the final sample. The athletes had been members of their respective teams for 31.57 ± 20.97 months. From an individual athlete perspective, female athletes comprised 46.4% of the sample (mean age = 21.09 ± 3.35); male athletes, 53.6% (mean age = 20.77 ± 2.75).

The teams competed predominantly at the university-level ($n = 147$) with a substantially smaller number competing in clubs ($n = 42$), intramurals ($n = 2$), and professional sport ($n = 1$). The types of sports represented included ice hockey ($n = 41$), volleyball ($n = 33$), basketball ($n = 31$), tennis ($n = 17$), netball ($n = 15$), soccer ($n = 14$), cricket ($n = 13$), ringette ($n = 9$), cheerleading ($n = 5$), track and field ($n = 3$), dance, synchronized swimming, and rugby ($n = 2$ in each), and field hockey, rowing, swimming, precision skating, and wrestling ($n = 1$ in each).

To explore one of the issues under investigation, the various sports also were classified into one of two categories on the basis of the relative amount of task

interdependence required to successfully achieve the team's goals. The sports forming the *interdependent* category ($n = 159$ teams) included hockey, volleyball, basketball, netball, soccer, ringette, rugby, and field hockey. The sports forming the *coactive/independent* category ($n = 32$) included tennis, cricket, cheerleading, track and field, dance, synchronized swimming, precision skating, swimming, and wrestling. In the case of the one rowing team, it was not possible to ascertain after the fact whether the athletes had competed in singles events or as a crew. Consequently, these data were not used in the analyses.

Measure

As indicated above, the Group Environment Questionnaire (GEQ; Carron et al., 1985) was used to assess cohesion. The GEQ is a self-report questionnaire that contains 18 items. Four aspects of cohesion are assessed: Individual attractions to the group-task (ATG-T, 4 items), Individual attractions to the group-social (ATG-S, 5 items), Group integration-task (GI-T, 5 items) and Group integration-social (GI-s, 4 items). Responses are obtained on a 9-point Likert-type scale anchored at the extremes by “*strongly disagree*” (1) and “*strongly agree*” (9). Thus, larger scores reflect stronger perceptions of cohesiveness. The Cronbach alpha values computed with the sample used in the present study were ATG-S ($\alpha = 0.610$), ATG-T ($\alpha = 0.715$), GI-T ($\alpha = 0.715$) and GI-S ($\alpha = 0.761$).

The alpha value for ATG-S was slightly lower than would typically be considered acceptable for research purposes (Nunnally, 1978). However, as was pointed out above, consensus (agreement) and reliability (internal consistency) are independent constructs (Kozlowski & Hattrup, 1992). It is possible to have high internal consistency and little consensus or vice versa. Because we were interested in consensus in the

present study, the ATG-S was retained in the analyses.

Procedures: Data collection

Although the data were obtained from a number of individual research projects, a series of general principles were adhered to in each study. These principles are part of the protocol used to collect cohesion data in our laboratories and they are briefly outlined below.

Coaches were contacted, the purpose of the study explained, and permission was sought to approach athletes for their participation in the study. When the coach's permission was secured, the investigators met with the athletes to explain the study and request their participation. In every instance, individual participation was voluntary and informed consent was obtained from all participants. In cases where individuals were under the age of 18, parental consent also was obtained.

Questionnaires were administered at a convenient team meeting or practice that was neither immediately before nor after a competition in order to avoid competition-specific biases in responses. Although all members of each team completed the questionnaire at a group meeting, the importance of independent responses was stressed in instructions at the time of administration. Participants were required to complete their questionnaires on their own and without conversation with their teammates.

Procedure: Statistical Analyses

Nonadjusted index of agreement. The index of agreement ($r_{wg(j)}$) reflects the proportion of systematic variance resulting from common judgments of a target stimulus in relation to the total variance (James et al., 1984). Consistent with the protocol outlined by James (1982) and James et al. (1984), the average item variance was used to represent the systematic variance (e.g., in the case of ATG-T, this was the average

variance from the four items in the scale). The total variance—which is referred to hereafter as the *expected variance*—also was computed using formulae outlined by James et al. (1984).

The formula used to compute the expected variance is based on the assumption that all options of the measurement scale (e.g., all 9 choices available to participants responding to GEQ items) have an equal likelihood of being chosen. As a consequence, the expected variance is also referred to as a *rectangular* or *null distribution*.

Adjusted index of agreement. James et al. (1984) also pointed out that there are instances where the assumption of a rectangular distribution might be unwarranted. That is, systematic bias might be present in the responses simply because of response set, social desirability, or positive leniency. Given that the index of agreement takes into account the amount of systematic variance in relation to the total (expected) variance, any systematic bias present because of positive leniency, for example, could lead to an inflated index of agreement. Consequently, an adjusted index of agreement is computed taking into account the nature of the biased distribution.

Typically, the nature and degree of bias are estimated--the data set under investigation is not used to determine if the distribution of responses is biased (skewed). However, the present data set was obtained from (a) an exceptionally large number of individual respondents, (b) an exceptionally large number and variety of teams and (c) a large number of investigations and investigators. Thus, rather than estimate the amount of bias, the data were used and the proportions of responses to each of the 9-choices for all of the four scales of the GEQ were computed. The results are summarized in Table 1. It is evident that the distributions for each of the four scales are positively skewed with athletes showing an increased likelihood of responding at the top half of the scale. Therefore, using protocol outlined by James et al. (1984), and the proportions presented

in Table 1, an adjusted index of agreement also was computed that took into account the statistical bias in the expected variance. It is worth noting that the “true” index of agreement is generally considered to fall in the range between the adjusted and nonadjusted values (James et al., 1984).

Out of range values. The expected variance is a theoretical (computed) value that is based on the assumption of a random selection of responses to the 9-point scale (i.e., thereby producing the so-called rectangular distribution). As a consequence, the expected variance is typically greater than the computed systematic variance. On occasion, however, members of a group could exhibit considerable disagreement about the target stimulus (e.g., the cohesiveness of the group). If this was the case, the systematic variance would turn out to be greater than the expected variance. When this occurs, $r_{wg(j)}$ assumes values outside the range of .00 to 1.00. James et al. (1984) have suggested that in these instances, the index of agreement for the group should be set to .00.

Results

Descriptive Statistics

Table 2 contains a summary of the nonadjusted and adjusted index of agreement values. It is apparent that the index of agreement varies widely between teams (i.e., range from $r_{wg(j)} = 0$ to .970) and across the four types of cohesion. As Table 2 shows, GI-T and GI-S were the cohesion constructs showing the greatest amount of shared beliefs among team members, especially for the nonadjusted values; ATG-S was the cohesion construct where shared beliefs were the lowest. Also, if $r_{wg(j)} = .50$ is used as a cut-off point, a greater percentage of teams showed lower agreement for ATG-T and ATG-S than for GI-T and GI-S. Typically, a greater percentage of teams exhibited a

serious lack of consensus (i.e., $r_{wg(j)} = .00$) for ATG-T and ATG-S than for GI-T and GI-S.

One question that does arise is whether the nonadjusted or adjusted index of agreement values better represent the extent to which shared beliefs about cohesion are present. One possible answer lies in how the data in Table 1 are interpreted. James et al. (1984) cautioned that the presence of a skewed distribution such as is evident in Table 1 does not automatically indicate the existence of response bias, positive leniency, or social desirability. If a target stimulus, an athlete's performance for example, is clearly outstanding, all judges would/should rate it uniformly high. In this example, correcting for response bias would lead to an underestimation in the amount of agreement present.

Similarly, the data in Table 1 may not reflect positive leniencies or response biases for cohesiveness—they may reflect the actual state of the situation being judged. As Donnelly et al. (1978) suggested, a non-cohesive group is literally a contradiction in terms. If a group continues to exist, it must be cohesive to some extent. Athletes actively and chronically involved in matters of importance to their teams might not use the full range of the 9-point cohesion scale because the lowest values (i.e., 1, 2, 3) simply would not be applicable. Further, disenchanted team members who might be inclined to perceive their group in a highly negative light are likely to withdraw from membership. If the tenure of group membership of the present sample is any indication, the athletes were not disenchanted team members. The average tenure at the time of testing was 31½ months. Also, approximately 75% of the sample had been members of their team for over 12 months. Thus, it is reasonable to assume that the athletes likely valued their membership because they maintained their involvement over a number of years.

There is no empirical way to determine if the skewness in response on the GEQ (presented in Table 1) is a product of positive leniency, social desirability, response

distortion, or other factors. However, the nature of the cohesion construct coupled with the duration of team membership of the athletes lead us to believe that the data in Table 1 most likely reflect the athletes' objective perception of the nature of their team's cohesiveness. As a consequence, we also believe that the nonadjusted index of agreement values reported here best reflect the extent of shared beliefs in cohesiveness in sport teams. Both unadjusted and adjusted statistics are presented to illustrate the range of possible values.

Type of Cohesion

In order to compare the magnitude of the nonadjusted indexes of agreement for the four measures of cohesion, a one-way MANOVA was computed. The Wilks' Lambda was significant, $F(3, 189) = 18.72, p < .001; \eta^2 = .23$. Post hoc Bonferroni analyses showed that the index of agreement values for GI-T and GI-S were significantly larger than for both ATG-S and ATG-T ($p < .03$). Further, the nonadjusted index of agreement values for GI-T and GI-S were not significantly different from each other; nor were those for ATG-T and ATG-S. Thus, the hypothesis that the strongest shared beliefs would be evidenced with GI-T and the weakest would be with ATG-S was supported for the nonadjusted index of agreement values.

Different results were obtained when the adjusted index of agreement values were analysed. Although the overall Wilks Lambda was again significant, $F(3,198) = 6.21, p < .001, \eta^2 = .09$, post hoc Bonferroni analyses showed that the adjusted index of agreement value for GI-S was significantly greater than the values for both ATG-S ($p < .001$) and GI-T ($p < .05$); there was no significant difference between GI-S and ATG-T. Also, the adjusted index of agreement values for GI-T, ATG-T, and ATG-S did not differ from one another.

Nature of the Task

Table 3 presents an overview of the index of agreement values for categories of teams classified on the basis of task interdependence. Generally, teams that are coactive/independent showed slightly larger nonadjusted and adjusted index of agreement values. However, MANOVA analyses with task type as the independent variable and the index of agreement for the four types of cohesion as the dependent variable showed that no significant differences ($p > .05$) were present.

Absolute Level of Cohesion in the Group

In order to determine whether the absolute level of cohesiveness in a team influences the relative degree of consensus, a tertile split of the total sample ($N= 192$ teams) was carried out independently with each of the four cohesion scales. For all four cohesion scales, the tertile split resulted in three categories of teams that differed significantly in amount of cohesion present ($p < .001$). An overview of the cohesion scores and index of agreement values is provided in Table 4.

As Table 4 shows, for the four GEQ scales the pattern of results was identical—the nonadjusted index of agreement increased as the absolute level of cohesiveness in the team increased. A series of one-way ANOVA were computed with categories of teams varying in cohesion representing the independent variable and the nonadjusted index of agreement representing the dependent variable. In all four one-way ANOVA, significant difference ($p < .001$) was observed among the three categories of teams.

Bonferroni post hoc analyses yielded a relatively consistent pattern of differences. That is, for all four cohesion scales the teams possessing the highest and intermediate levels of cohesiveness had significantly (p at least $.05$) higher index of agreement values than teams possessing the lowest levels of cohesiveness. More specifically, for ATG-T and ATG-S only, the teams possessing the highest amounts of

cohesiveness also had significantly (p at least .05) greater index of agreement values than teams possessing the lowest levels of cohesiveness.

An examination of the adjusted index of agreement values in Table 4 shows similar linear pattern to the above. That is, the adjusted index of agreement values show an increase as cohesion increases. When a series of one-way ANOVA were computed, again a significant difference ($p < .001$) was observed among the three categories of teams.

Bonferroni post hoc analyses showed the identical pattern of significant results obtained for the nonadjusted index of agreement values. That is, for all four cohesion scales the teams possessing the highest and intermediate levels of cohesiveness had significantly (p at least .05) higher index of adjusted agreement values than teams possessing the lowest levels of cohesiveness. In the case of ATG-T and ATG-S only, the teams possessing the highest amounts of cohesiveness also had significantly (p at least .05) greater adjusted index of agreement values than teams possessing the lowest levels of cohesiveness. These results failed to show support for the hypothesis that a curvilinear relationship between team cohesion levels and the index of agreement would be present (i.e., that consensus would be greater when team cohesion was high and low versus when team cohesion was at a moderate/intermediate level). The results did show evidence for a linear relationship; consensus increases as team cohesiveness increases.

Discussion

The general purpose of the study was to examine the range and absolute magnitude of shared beliefs--the index of agreement ($r_{wg(j)}$)—present for group cohesion in sport teams. Within the framework of this general purpose, there were three secondary purposes--to examine how the type of cohesion, the task interactive nature of the group,

and the absolute level of cohesion present are related to the index of agreement exhibited for team cohesiveness.

The results showed that the average nonadjusted index of agreement values varied from $r_{wg(j)} = .563$ (ATG-S) to $.721$ (GI-T) and the adjusted index of agreement values from $r_{wg(j)} = .449$ (ATG-S) to $.558$ (GI-S). With two exceptions, the median index of agreement values were above $r_{wg(j)} = .600$ and with one exception, below $r_{wg(j)} = .800$.

One question that arises is what absolute magnitude of $r_{wg(j)}$ should researchers and practitioners use to conclude that groupness—a shared belief about cohesiveness—is present in any given team? Based upon classical test theory (with collective efficacy as their prototype for shared belief), Moritz and Watson (1998) proposed that $r_{wg(j)}$ values in the range of $.50$ to $.80$ are likely sufficient for research purposes, but values above $.90$ may be necessary for applied purposes. The latter value seems unreasonably high for the construct of cohesion as measured by the GEQ. As Dion has pointed out, conceptually cohesion is a property of the group but it is *manifested at the individual level* in attractions to the group, to its members, to its tasks and goals. As Table 2 shows, if nonadjusted $r_{wg(j)}$ values above $.50$ are used as a criterion, 35.4%, 26.0%, 14.1%, and 15.6% of the teams tested would be considered to have insufficient agreement in terms of ATG-S, ATG-T, GIS, and GI-T respectively to be considered a group. When we know so little about the agreement characterizing the shared beliefs of real-world groups, and we consider the number of factors (e.g., nature of measurement scales, Hinsz et al, 1997) that can influence degree of agreement, it may be wise to consider different conventions for different research decisions. At this early stage of research on cohesion-as-shared beliefs, the danger of making a Type II error by discarding data seems probable.

What $r_{wg(j)}$ values should be used in order to conclude that sufficient consensus is present and team members have shared beliefs relative to the various manifestations of

cohesiveness? Using the current sample and the nature of the GEQ as a basis, we propose that small, moderate and large degrees of groupness in perceptions of cohesion can be assumed to be present when nonadjusted $r_{wg(j)}$ values of .40, .50 and .60 are present for ATG-T and ATG-S and nonadjusted $r_{wg(j)}$ values of .50, .60 and .70 are present for GI-T and GI-S.

The different standards proposed for the Individual attractions to the group (ATG-S and ATG-T) versus the Group integration (GI-S and GI-T) manifestations of cohesiveness are based on the fact that the athlete's focus is different for each (i.e., "me" in the case of the Individual attractions to the group subscales versus the "the group" in the case of the Group integration subscales). Therefore, the amount of consensus that reasonably could be expected should also differ. Klein, Conn, Smith, and Sorra (2001) pointed out that within-group consensus is strongly influenced by focus; that is, by where respondents are asked to direct their attention. Experiences or situations requiring "I", "my", and/or "me" evaluations encourage respondents to focus within themselves and disregard the experiences of other group members. Conversely, experiences or situations requiring "we", "us", and/or "the team" evaluations encourage respondents to focus outside themselves and consider common experiences as a reference. Agreement should be higher in the latter than in the former case.

The recommended index of agreement values for ATG-S and ATG-T may seem low. However, it should be reemphasized that conceptually both ATG-S and ATG-T are considered to reflect a *group property*—why athletes cohere and stay united in pursuit of instrumental objectives and to satisfy affective needs (Carron et al., 1998). If practical or research-oriented situations arise where aggregation of any of the four manifestations of cohesiveness is meaningful, removing teams from consideration because they fail to meet unreasonably high index of agreement standards is counterproductive.

In general, most group studies do not have the size of the database examined in the present study. Most samples are considerably smaller and statistical analyses using the group as the unit of analysis may suffer a loss of power if conventions from classical test theory are used arbitrarily. Instead, we suggest that the data, the type of construct, and the nature of the research question should determine the selection of an agreement standard.

The three analyses associated with the secondary purposes of the study illustrate that the conditions under which individual consensus about cohesiveness is achieved within teams is complex. Consensus is more likely when athletes evaluate the collective unity around the task objectives (i.e., GI-T) and least likely when they evaluate how the team satisfies their own personal social needs (i.e., ATG-S). These results support the Hinsz et al. (1997) contention that, “by their nature, groups are context sensitive and context situated” (p. 45). Intercollegiate and club-level sport teams are inherently task-oriented; group success is a major priority for all teams. As a consequence, information associated with the team’s common task is more likely to be universally shared within the team than information about how the group satisfies each athlete’s personal social needs.

Surprisingly, the results showed that the amount of interaction inherent in the nature of the sport is not related to the level of agreement about cohesiveness. Within organizational psychology, a substantial body of research has shown that the amount of interaction in the group—either task or social—is positively related to increased similarity in perceptions, cognitions, and behavior (e.g., Klein et al., 2001; Salancik & Pfeffer, 1978; Walsh & Ungson, 1991; Weick, 1979). Relative to the number of teams in the interactive category ($n = 159$), there were few teams in the coactive/independent category ($n = 32$). This may have been one factor in the failure to show differences. Also, however, perhaps the present results draw into question the assumption that there are

more task- and socially-related interactions that promote consensus in sports such as basketball than in sports such as track and field. Alternatively, the present results may draw into question the assumption that the greater number of task- and socially-related interactions in sports such as basketball ipso facto lead to greater consensus about the team's cohesiveness.

The results also showed that the absolute level of cohesion present is strongly related to the average amount of consensus present. Team members are more likely to show higher consensus about their team's cohesiveness when perceive that cohesiveness to be greater. Despite the fact that a curvilinear relationship between consensus and cohesiveness was proposed, in hindsight our results may not be surprising. It seems improbable that high consensus (i.e., shared agreement) about *low cohesion* could be measured while the team was still together. An exception would be unique conditions that force a group together, yet encourage shared agreement on lack of cohesion. If the psychological nature of the situation is sufficiently powerful, the wide individual differences in individual cognitions, behaviors, and affect are reduced and consistency across individuals results (Mischel, 1973). Thus, consensus and level of cohesion may be a unidirectional relationship in contrast to the relationship between member agreement and collective efficacy (i.e., strong agreement about either high or low collective efficacy) suggested by Moritz and Watson (1998).

The present results are interesting in light of Moritz and Watson's (1998) response to their own question about what researchers should do if teams show no consensus:

Guzzo, Yost, Campbell and Shea (1993) stated that low agreement may indicate that the group is not a consequential entity to its members or that the unity of the group is doubtful. Thus, a researcher might properly

choose to exclude from further analyses those groups having insufficient intermember agreement” (p. 296)

There are at least two reasons why a decision to discard teams from further analyses solely on the basis of a low index of agreement could be unsound scientifically. The first reason is statistical; discarding teams leads to a loss of statistical power and an increase in the probability of committing a Type II error. The second, related but more important reason is conceptual. Some research questions require the inclusion/retention of all groups—not just those groups who exhibit more “groupness” as evidenced through their strong consensus on the construct of interest. A good example is the question of the relationship of cohesion to team success in sport. Team success is not measured at the individual level; an aggregated response (i.e., a team’s winning percentage) must be used. Historically, researchers testing the relationship of team cohesion to team success have not shown interest in the question, “what is the relationship of cohesion to performance in those teams where members exhibit relatively good agreement on the degree to which their team is cohesive?”

A recent meta-analysis has shown unequivocally that higher group cohesiveness in all of its manifestations—ATG-S, ATG-T, GI-S, and GI-T—is associated with greater team success (Carron, Colman, Wheeler & Stevens, 2002). If researchers contributing to the meta-analysis had used the criterion of $r_{wg(j)} = .50$ to reflect “groupness” and support aggregation (and our data can be assumed to reflect population values), 35.5%, 26.0%, 14.1%, and 15.6% of their sample would have been discarded for analyses of the relationship between performance and ATG-S, ATG-T, GI-S, and GI-T respectively. It seems likely that the reduction in team differences in cohesion (and increased homogeneity in the sample) would have contributed to a dramatically different picture of the cohesion-performance relationship in sport

In conclusion, a point that is important from the above discussion is that research decisions are based on a myriad of factors—some empirical, some conceptual. There is no gold standard available for researchers insofar as acceptable sample sizes, probability levels, and effect sizes. Questions concerning what is an acceptable index of agreement should be considered in a similar light.

The present study is largely a preliminary, descriptive examination of the nature of the shared beliefs that reflect several aspects of cohesion in sport groups. To the best of our knowledge, the study represents the first statistical examinations of the level of shared agreement in cohesive groups, and as such, addresses part of Dion's (2000) call for an examination of the index of agreement as part of the practice necessary for aggregation of individual member responses to the group level. As Dion reminded us, "one perspective is that "cohesiveness ... means different things for different groups with different goals and tasks" (p.22). Thus, It would be informative if researchers continue to explore the notion of shared beliefs and search out the group factors that moderate and mediate shared beliefs in cohesiveness in sport teams and their relation to both individual and/or group outcomes.

References

- Bernthal, P. R., & Insko, C. A. (1993). Cohesiveness without groupthink. *Group & Organization Management, 18*, 66-87.
- Bollen, K. A., & Hoyle, R. H. (1990). Perceived cohesion: A conceptual and empirical examination. *Social Forces, 69*, 479-504.
- Brawley, L.R., Carron, A.V. & Widmeyer, W. N. (1987). Assessing the cohesion of sport teams: Validity of the Group Environment questionnaire. *Journal of Sport Psychology, 9*, 275-294.
- Brawley, L. R., Carron, A. V. & Widmeyer, W. N. (1988). Exploring the relationship between cohesion and group resistance to disruption. *Journal of Sport & Exercise Psychology, 10*, 199-213.
- Carron, A.V., Brawley, L.R., & Widmeyer, W.N. (1998). Measurement of cohesion in sport and exercise. In J. L. Duda (Ed.), *Advances in sport and exercise psychology measurement*, (pp. 213-226). Morgantown, WV: Fitness Information Technology.
- Carron, A. V., Bray, S. R., & Eys, M. A. (2002). Team cohesion and team success in sport. *Journal of Sport Sciences, 20*, 119-126.
- Carron, A. V., Colman, M. M., Wheeler, J., & Stevens, D. (2002). Cohesion and performance in sport: A meta-analysis. *Journal of Sport and Exercise Psychology, 24*, 168-188.
- Carron, A.V. Widmeyer, W.N., & Brawley, W. N. (1985). The development of an instrument to measure cohesion in sport teams: The Group Environment Questionnaire. *Journal of Sport Psychology, 7*, 244-266.

- Carron, A.V., Widmeyer, W.N., & Brawley, L.R. (1988). Group cohesion and individual adherence to physical activity. *Journal of Sport and Exercise Psychology, 10*, 127-138.
- Cole, M. (1991). Conclusion. In L. B. Resnick, J. M. Levine, & S. D. Teasley (Eds.), *Perspectives on socially shared cognition*. (pp. 399-417). Washington, DC: American Psychological Association.
- Dion, K. R. (2000). Group cohesion: From “field of forces” to multidimensional construct”. *Group Dynamics: Theory, Research and Practice, 4*, 7-26.
- Donnelly, P., Carron, A.V., & Chelladurai, P. (1978). *Group cohesion and sport*. Ottawa, Ont.: Canadian Association for Health, Physical Education, & Recreation Monograph Series.
- Festinger, L., Schachter, S., & Back, K. (1950). *Social presses in informal groups: A study of a housing project*. New York: Harper.
- Gross, N., & Martin, W. (1952). On group cohesiveness. *American Journal of Sociology, 57*, 533-546.
- Guzzo, R. A., Yost, P. R., Campbell, R. J., & Shea, G. P. (1993). Potency in groups: Articulating a construct. *British Journal of Social Psychology, 32*, 87-106.
- Hinsz, V. B., Tindale, R. S., & Vollrath, D. A. (1997). The emerging conceptualisation of groups as information processors. *Psychological Bulletin, 121*, 43-64.
- Janis, I. L. (1982). *Groupthink: Psychological studies of policy decisions and fiascos*. Boston, MA: Houghton-Mifflin.
- James, L. R. (1982). Aggregating bias in estimates of perceptual agreement. *Journal of Applied Psychology, 67*, 219-229.

- James, L. R., Demaree, R. G., & Wolf, G. (1984). Estimating within-group interrater reliability with and without response bias. *Journal of Applied Psychology, 69*, 85-98.
- Kahneman, D. (1973). *Attention and effort*. Englewood Cliffs, NJ: Prentice-Hall.
- Klein, K. J., Conn, A. B., Smith, D. B., & Sorra, J. S. (2001). Is everyone in agreement? An exploration of within-group agreement in employee perceptions of the work environment. *Journal of Applied Psychology, 86*, 3-16.
- Kozlowski, S. W. J., & Hattrup, K. (1992). A disagreement about within-group agreement: Disentangling issues of consistency versus consensus. *Journal of Applied Psychology, 77*, 161-167.
- Levine, J. M., Resnick, L. B., & Higgins, E. T. (1993). Social foundations of cognition. *Annual Review of Psychology, 44*, 585-612.
- Mischel, W. (1973). Toward a cognitive social learning reconceptualization of personality. *Psychological Review, 80*, 252-283.
- Moritz, S. E., & Watson, C. B. (1998). Levels of analysis issues in group psychology: Using efficacy as an example of a multilevel model. *Group Dynamics: Theory, Research, and Practice, 2*, 285-298.
- Nunnally, J. C. (1978). *Psychometric theory*. New York: McGraw-Hill.
- Salancik, G. R., & Pfeffer, J. (1978). A social information processing approach to job attitudes and task design. *Administrative Science Quarterly, 23*, 224-253.
- Schmidt, F. L., & Hunter, J. E. (1989). Interrater reliability coefficients cannot be computed when only one stimulus is rated. *Journal of Applied Psychology, 74*, 368-370.
- Treisman, A. (1988). Features and objects: The Fourteenth Bartlett Memorial Lecture. *Quarterly Journal of Experimental Psychology, 40A*, 401-407.

Treisman, A., & Gelade, G. (1980). A feature integration theory of attention. *Cognitive Psychology, 12*, 97-136.

Walsh, J. P., & Ungson, G. R. (1991). Organizational memory. *Academy of Management Review, 16*, 57-91.

Weick, K. E. (1979). *The social psychology of organizing*. Reading, MA: Addison Wesley.

Widmeyer, W.N., Brawley, L.R. & Carron, A.V. (1985). *Measurement of cohesion in sport teams: The Group Environment Questionnaire*. London, Ont.: Spodym Publishers.

Footnote

1. Exceptions to this general rule were studies carried out independently by former students and/or collaborators of the first two investigators.

Table 1.

Proportion of Responses to Each of the 9 Choices Available to Respondents on the Group Environment Questionnaire

	1	2	3	4	5	6	7	8	9
ATGS	.049	.043	.052	.042	.094	.063	.127	.173	.355
ATGT	.031	.036	.055	.058	.077	.064	.124	.201	.396
GIS	.038	.061	.091	.086	.141	.092	.162	.176	.154
GIT	.027	.037	.063	.071	.121	.101	.182	.209	.187

(n = 2107)

Table 2.

*Descriptive Statistics for the Index of Agreement ($r_{(wg(j))}$) for the Scales of the Group**Environment Questionnaire (n = 192 teams)*

Variable ¹	Mean \pm S.D. ²	Median	Percentage of teams with $r_{(wg(j))}$ values		
			= .00	< .50	> .80
Nonadjusted	.563 \pm .273	.625	7.8	35.4	22.9
ATG-S					
Adjusted ATG-S	.449 \pm .325	.508	22.9	49.5	17.2
Nonadjusted	.621 \pm .275	.686	8.9	26.0	32.2
ATG-T					
Adjusted ATG-T	.496 \pm .324	.557	17.7	42.7	22.4
Nonadjusted GI-S	.694 \pm .217	.763	2.1	14.1	41.6
Adjusted GI-S	.558 \pm .294	.658	11.5	32.8	23.4
Nonadjusted GI-T	.721 \pm .236	.806	3.6	15.6	53.6
Adjusted GI-T	.489 \pm .350	.606	26.0	42.2	22.9

1. ATG-T = Individual attractions to the group-task; ATG-S = Individual attractions to the group-social; GI-T = Group integration-task; GI-S = Group integration-social.

2. Responses on the Group Environment Questionnaire are provided on a 9-point scale; higher responses reflect greater cohesiveness.

Table 3.

Index of Agreement for Interactive (n = 159) and Coactive/Independent (n = 32) Teams

Variable	Interactive	Coactive/Independent	p
	Teams.	Teams	
Nonadjusted ATG-S	.561 ± .27	.565 ± .31	ns
Adjusted ATG-S	.439 ± .32	.485 ± .32	ns
Nonadjusted ATG-T	.590 ± .29	.770 ± .14	ns
Adjusted ATG-T	.459 ± .33	.673 ± .22	ns
Nonadjusted GI-S	.703 ± .21	.658 ± .24	ns
Adjusted GI-S	.573 ± .28	.497 ± .35	ns
Nonadjusted GI-T	.718 ± .23	.726 ± .25	ns
Adjusted GI-T	.478 ± .34	.531 ± .35	ns

ATG-T = Individual attractions to the group-task; ATG-S = Individual attractions to the group-social; GI-T = Group integration-task; GI-S = Group integration-social

Table 4.

Index of Agreement Values for Groups Differing in their Absolute Amount of Cohesion.

Group	Number of Teams	Average Cohesion ¹	Nondjusted Index of Agreement	Adjusted Index of Agreement
Individual Attractions to the Group-Task				
Low Cohesion	64	5.83±.53	.422	.264
Moderate Cohesion	63	6.97±.27	.612	.453
High Cohesion	65	7.90±.30	.825	.764
Individual Attractions to the Group-Social				
Low Cohesion	64	5.93±.45	.415	.278
Moderate Cohesion	65	6.92±.18	.581	.462
High Cohesion	63	7.59±.30	.695	.609
Group Integration-Social				
Low Cohesion	63	4.61±.60	.600	.466
Moderate Cohesion	64	5.96±.36	.654	.489
High Cohesion	65	7.32±.43	.826	.751

Cohesion				
Group Integration-Task				
Low Cohesion	64	5.44 \pm .48	.642	.356
Moderate Cohesion	65	6.48 \pm .21	.690	.438
High Cohesion	63	7.31 \pm .24	.829	.670

1. Responses on the cohesion scales could vary from 1 (low cohesiveness) to 9 (high cohesiveness)