# Small Group Research

http://sgr.sagepub.com

## **Cohesion as Shared Beliefs in Exercise Classes**

Shauna M. Burke, Albert V. Carron, Michelle M. Patterson, Paul A. Estabrooks, Jennie L. Hill, Todd M. Loughead, Sara R. Rosenkranz and Kevin S. Spink Small Group Research 2005; 36; 267 DOI: 10.1177/1046496404267942

The online version of this article can be found at: http://sgr.sagepub.com/cgi/content/abstract/36/3/267

## Published by: SAGE http://www.sagepublications.com

#### Additional services and information for Small Group Research can be found at:

Email Alerts: http://sgr.sagepub.com/cgi/alerts

Subscriptions: http://sgr.sagepub.com/subscriptions

Reprints: http://www.sagepub.com/journalsReprints.nav

Permissions: http://www.sagepub.com/journalsPermissions.nav

Citations http://sgr.sagepub.com/cgi/content/refs/36/3/267

# COHESION AS SHARED BELIEFS IN EXERCISE CLASSES

SHAUNA M. BURKE ALBERT V. CARRON MICHELLE M. PATTERSON University of Western Ontario PAUL A. ESTABROOKS JENNIE L. HILL Kaiser Permanente Colorado TODD M. LOUGHEAD McGill University SARA R. ROSENKRANZ USA Triathalon KEVIN S. SPINK

**KEVIN S. SPINK** University of Saskatchewan

The purpose of the study was to determine if perceptions of cohesion in exercise classes demonstrated sufficiently high consensus and between-group variance to support a conclusion that exercise classes are groups. Participants (N = 1,700) in 130 classes were tested on either the Group Environment Questionnaire (GEQ) or the Physical Activity Group Environment Questionnaire (PAGEQ). Results showed that exercise classes satisfied the statistical criteria necessary to support a conclusion that they are true groups; that is, they exhibited acceptable levels of consensus about cohesion within classes and acceptable differences in cohesion between classes. In addition, index-of-agreement values were significantly greater for participants completing the PAGEQ than for participants completing the GEQ. Finally, consensus was greatest when participants evaluated how the exercise class satisfied their own personal task needs (i.e., individual attractions to the group-task), and second greatest when participants evaluated the collective unity around the task objectives (i.e., group integration-task).

Keywords: consensus; index of agreement; Group Environment Questionnaire; Physical Activity Group Environment Questionnaire

SMALL GROUP RESEARCH, Vol. 36 No. 3, June 2005 267-288 DOI: 10.1177/1046496404267942 © 2005 Sage Publications

267

The general focus of the current study was to determine whether perceptions of cohesion in exercise classes demonstrated sufficiently high consensus and between-group variance to support a conclusion that exercise classes are groups. Over time, several definitions have been proposed in efforts to provide a clear and precise definition that can be used to examine the nature, antecedents, and consequences of various group processes. In 1950, Festinger, Schachter, and Back defined a group as "a number of interacting and sociometrically connected people" (p. 58). McGrath (1984) defined groups as "social aggregates that involve mutual awareness and potential mutual interaction" (p. 7). McGrath argued that although most social aggregates consist of two or more people, they are not necessarily groups-in fact, several social aggregations do not involve the potential for mutual interaction and therefore cannot be considered true groups. For instance, artificial aggregations (i.e., statistical groups or social categories formed based on factors such as social class, age, or sex) and units with patterned relationships (i.e., cultures, subcultures) are not classified as true groups (McGrath, 1984). In addition, unorganized aggregates (e.g., crowd, audience, public), structured social units (e.g., society, community), deliberately designed social units (e.g., organizations, suborganizations) and less deliberately designed social units (e.g., associations) do not qualify as groups based on McGrath's classification because of the fact that mutual awareness and interaction are not necessarily present.

Additional criteria incorporated into various definitions of groups have included common fate, mutual benefit, social structure, group processes, and self-categorization (Carron & Hausenblas, 1998). The notion of common fate suggests that regardless of individual actions, group events and group outcomes affect all group members. For example, some athletes may contribute more than their

AUTHORS' NOTE: Address correspondence concerning this article to Shauna M. Burke, School of Kinesiology, University of Western Ontario, London, Ontario, N6A 3K7, Canada; e-mail: sburke7@uwo.ca

teammates in competition; however, it is the entire team that wins or loses. Mutual benefit for group members speaks to the idea that belonging to a group is rewarding for individual members, whereas sitting among a crowd of people in a waiting room does not yield the same benefits. Another common way to characterize a group is through its social structure. As a result of group member relationships, true groups evolve in structure and come to possess group norms, roles, and status differences (Carron & Hausenblas, 1998). For obvious reasons, group processes such as communication and group member relations are also included in definitions of groups. It has been argued that a collection of individuals attending a concert might interact or communicate with one another to some extent. Therefore, many authors refer to the quality of relationships among group members in their definitions (Carron & Hausenblas, 1998). Last, groups are characterized by self-categorization. Selfcategorization refers to individual group members' perceptions and feelings of belonging to an actual group.

On the one hand, a considerable amount of research has shown that perceptions of group cohesion in exercise classes are reliably related to several important outcomes. Specifically, during the past 25 years, research has shown that individual behaviors (e.g., adherence), affect (e.g., satisfaction, attitude), and cognitions (e.g., perceptions of resistance to disruptive events) associated with exercise are positively influenced by group cohesion (Carron, Hausenblas, & Mack, 1996; Carron, Widmeyer, & Brawley, 1988; Spink & Carron, 1992, 1993, 1994).

Conversely, however, from a group dynamics perspective, exercise classes possess few of the characteristics commonly associated with true groups (e.g., McGrath, 1984). For example, although exercise class participants may experience a common fate when their class is cancelled or when the instructor arrives late, important elements of group structure (i.e., group norms, group roles, status differences), group processes (i.e., collective decision making, cooperative behavior), and group products (e.g., group productivity and/or group success) are usually minimally present or absent altogether (Carron & Hausenblas, 1998).

#### EMPIRICAL CRITERIA FOR THE EXISTENCE OF A GROUP

One indicant of the degree to which collections of individuals are interdependent in a meaningful way is the presence of shared beliefs (Moritz & Watson, 1998). Empirically, the degree to which shared beliefs are present can be determined with the index of agreement,  $r_{wg(j)}$  (James, 1982; James, Demaree, & Wolf, 1984; Kozlowski & Hattrup, 1992; Schmidt & Hunter, 1989). In essence, the index of agreement,  $r_{wg(j)}$ , provides a statistical measure of the extent to which members of a unit (e.g., an exercise class) show consensus in their perceptions of a target stimulus (e.g., their class's level of cohesiveness). Mathematically,  $r_{wg(j)}$  is defined as "the proportion of systematic variance in a set of judgments in relation to the total variance in the judgments" (James et al., 1984, p. 86).

However, as Moritz and Watson (1998) pointed out, the existence of shared beliefs within collections of individuals-as reflected in relatively high index-of-agreement values—is not sufficient to conclude that aggregates of individuals are groups. It is possible, for independent, random collections of individuals to possess similarly high degrees of consensus about an issue of common importance. For example, collections of individuals at different locations along a bus route could be expected to demonstrate relatively high consensus on the bus's schedule. Thus, a second statistical criterion is used that involves contrasting the proportion of within- and between-group variance via one-way ANOVA with groups (i.e., exercise classes) as the independent variable and the construct of interest (i.e., cohesion) as the dependent variable (Kenny & Lavoie, 1985). In addition, to determine the relative strength of the between-groups effect (i.e., information not available though a F value by itself), the intraclass correlation coefficient (ICC) and etasquared statistics are also computed. Demonstrating a substantial proportion of between-group variability provides further support to conclude that a group phenomenon is present.

The main purpose of the current study was to examine whether perceptions of cohesion in exercise classes demonstrate sufficiently high consensus (as reflected in the index of agreement) and between-group variance (as reflected in the *F* value, the ICC, and the eta-squared) to support a conclusion that they are groups.

#### HYPOTHESES ASSOCIATED WITH WITHIN-CLASS CONSENSUS

It was hypothesized that the index of agreement—although values could vary across the four dimensions of cohesion examined—would be in a moderate range (i.e.,  $r_{wg[i]}$  from 0.40 to 0.60).<sup>1</sup>

The dimensions of cohesion examined—the operational definition-in the current study evolved from a conceptual model advanced by Carron and his colleagues (Carron, Brawley, & Widmeyer, 1998; Carron, Widmeyer, & Brawley, 1985). The basis of the conceptual model is that cohesiveness is reflected in 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 common experiences and are assessed by questions using plural nouns such as we. 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 idiosyncratic experiences and are assessed by singular nouns such as I. Carron and colleagues (Carron, Brawley, & Widmeyer, 1998; Carron, Widmeyer, & Brawley, 1985) also proposed that there are two fundamental orientations in a group member's perceptions: task and social. Thus, four manifestations of cohesion are proposed: 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).

Klein, Conn, Smith, and Sorra (2001) proposed that consensus should be higher when individuals are forced to focus on situations or experiences that require *we* or *us* evaluations (e.g., the group integration dimensions of cohesion) because they may encourage respondents to consider the shared perspective of the group as a whole. Conversely, statements containing *I*, *me*, or *my* evaluations (e.g., the individual attractions to the group dimensions of cohesion) require respondents to focus their attention on their personal experiences and to disregard the experiences of other group members. In addition, exercise classes are predominantly task oriented; that is, participants come to the group setting for primarily fitnessrelated reasons (see Carron, Hausenblas, & Estabrooks, 2003). Consequently, it was hypothesized that the dimension showing the strongest level of agreement would be GI-T and the dimension showing the weakest would be ATG-S. No hypothesis was advanced in respect to the relative level of agreement expected for GI-S and ATG-T.

#### HYPOTHESES ASSOCIATED WITH BETWEEN-GROUP VARIABILITY

It was also hypothesized that all dimensions of cohesion (i.e., ATG-T and ATG-S and GI-T and GI-S) would demonstrate sufficient between-group differences to support a conclusion that exercise classes are true groups. Support for this hypothesis comes from a substantial body of literature showing that exercise class participants' perceptions of cohesiveness are reliably related to various behaviors (e.g., adherence), cognitions (scheduling efficacy), and types of affect (e.g., satisfaction; see Carron, Hausenblas, & Estabrooks, 2003, for an overview).

#### METHOD

#### PARTICIPANTS

Data files from multiple investigations of cohesion in exercise classes as measured by the Group Environment Questionnaire (GEQ; Brawley, Carron, & Widmeyer, 1987; Carron, Widmeyer, & Brawley, 1985; Widmeyer, Brawley, & Carron, 1985) and the Physical Activity Group Environment Questionnaire (PAGEQ; Estabrooks & Carron, 2000) were combined to produce the data set (Carron & Spink, 1993, 1995; Estabrooks & Carron, 1999a and b; Hill & Estabrooks, 2000; Hill, Estabrooks, & Milliken, 2001; Loughead & Carron, 2004; Loughead, Colman, & Carron, 2002; Spink &

Carron, 1992, 1993, 1994). The data were from 1,700 exercise participants (mean age = 55.67, SD = 20.33) who had been members of 130 physical activity classes. Female exercise participants made up 83.7% of the sample and had a mean age of 55.35 (SD = 10.18). Male participants made up 16.3% of the sample and had a mean age of 62.85 (SD = 9.20). The average exercise class contained 12 participants, and the average age of the participants in the exercise classes ranged from 19.29 years to 84.38 years.<sup>2</sup>

#### MEASURES

Originally, the GEQ (Brawley et al., 1987; Carron, Widmeyer, & Brawley, 1985; Widmeyer et al., 1985) was developed to measure cohesion in sport. However, an exercise-modified version of the GEQ has been used to assess individual perceptions of cohesiveness in a number of the physical activity classes employed in the current sample (n = 31). The exercise-modified GEQ is an 18-item self-report questionnaire that assesses cohesion via perceptions of individual group members. Class members are required to respond to the 18 statements about their group on a 9-point Likert-type scale with anchors of *strongly disagree* and *strongly agree*. Thus, larger scores reveal stronger perceptions of cohesiveness among group members. As indicated above, the GEQ assesses four dimensions of cohesion: ATG-T (4 items), ATG-S (5 items), GI-T (5 items), and GI-S (4 items).

In the exercise-modified version of the GEQ, minor wording changes and changes in the situational frame of reference were made to ensure that participants were providing their perceptions of an exercise class rather than a sport team (Carron & Spink, 1992). Whereas an ATG-T item on the GEQ pertaining to sport states, "This team does not provide me with enough opportunities to improve my personal performance," the exercise-modified item read, "This exercise group does not give me enough opportunities to improve my physical fitness" (Spink & Carron, 1993).

The GEQ modified for the exercise domain has been shown to possess internal consistency values similar to the original version as well as adequate internal discriminant validities (Carron &

Brawley, 2000; Carron & Spink, 1992; Carron, Widmeyer, & Brawley, 1988; Spink & Carron, 1992, 1993, 1994). The Cronbach's alpha values for the current sample were ATG-T ( $\alpha = .748$ ), ATG-S ( $\alpha = .479$ ), GI-T ( $\alpha = .711$ ) and GI-S ( $\alpha = .709$ ). The alpha value for ATG-S was lower than the standard of  $\alpha = .70$  advocated by Nunnally (1978) for research purposes. However, Kozlowski and Hattrup (1992) pointed out that consensus (agreement) and reliability (internal consistency) are independent constructs—that it is possible to obtain low internal consistency and high consensus and vice versa. Because the current study focused on consensus rather than on reliability and it was the first investigation focusing on the issue of consensus in exercise classes, the ATG-S was retained for further analyses.

The more recently developed PAGEQ (Estabrooks & Carron, 2000) was used to assess individual perceptions of cohesiveness in the majority of the physical activity classes examined in the current study (n = 99). The PAGEQ is a 21-item self-report questionnaire that assesses individual member's perceptions of cohesion in relation to an exercise class. The four aspects of cohesion that are measured using the PAGEQ are identical to those measured by the GEQ. Thus, the four dimensions of the PAGEQ are ATG-T (6 items), ATG-S (6 items), GI-T (5 items), and GI-S (4 items).

With the PAGEQ, exercise participants are required to respond to the 21 statements about their group on a 9-point Likert-type scale ranging from 1 (*very strongly disagree*) to 9 (*very strongly agree*). Larger scores reflect stronger perceptions of cohesion. The Cronbach's alpha values computed for the PAGEQ in the current study were ATG-T ( $\alpha = .899$ ), ATG-S ( $\alpha = .910$ ), GI-T ( $\alpha = .754$ ) and GI-S ( $\alpha = .822$ ).

#### PROCEDURES

*Data collection*. Initially, the principle investigators of several studies that examined group cohesion in an exercise context were contacted. The rationale of the current study was explained, and permission was then sought for the use of the researchers' data sets.

In every instance, the researchers agreed to the use of their data for the purposes of the current study.

Although data were obtained from multiple investigations of cohesion in exercise classes, identical research principles were adhered to in each study. The principles that constitute the general protocol used to collect cohesion data are as follows.

Prior to class commencement, fitness class administrators and/ or instructors were individually contacted and the purpose of the study was outlined. Permission was then sought for the researcher(s) to administer the cohesion questionnaire to the exercise class participants at a specific time (e.g., at the completion of the 3rd week of classes). At this time, the most appropriate method for collecting data was determined, depending on the investigation (e.g., to minimize disruption to the exercise class; Spink & Carron, 1992). Data were collected during or after completion of the 3rd week of the session(s) to allow cohesion—a group property—sufficient time to develop (Spink & Carron, 1993).

The questionnaires and informed consent forms were distributed to each participant prior to the exercise session. In every study, participation was voluntary, and informed consent was obtained from all exercise participants. For those individuals younger than age 18 years, parental consent was obtained. In addition, participants were informed of the general purpose of the study and were assured that all responses would be completely confidential (cf. Spink & Carron, 1994). Questionnaires were filled out independently and anonymously, and research assistants and/or the study investigators collected the questionnaires from all participants on completion.

Statistical analyses. As indicated above, the index of agreement,  $r_{wg(j)}$ , was the statistic used to empirically estimate the level of agreement (i.e., consensus) present in exercise classes. The use of  $r_{wg(j)}$  in the current investigation was warranted for a number of reasons. First, contrary to other analyses, the index of agreement is sensitive to various degrees of agreement (James et al., 1984). Second, James and colleagues (1984) suggested that it is particularly

useful to assess within-group agreement for groups that consist of a number of respondents who rate a set of items that measure the same construct—in this case, cohesion as measured by the GEQ and the PAGEQ. It has also been argued that the index of agreement is highly functional when agreement is high or when the number of raters is large (James et al., 1984). In the current study, the average number of raters in each class (approximately 12 participants) was sufficient for use of this statistical technique.

James and colleagues (1984) recommended that two important conditions be met prior to assessing the level of agreement among raters. One condition is that there must be evidence that the scales used in the analyses have acceptable psychometric properties (i.e., reliability and construct validity). In prior research, the PAGEQ and the exercise-modified version of the GEQ have demonstrated adequate internal consistency and sound predictive, concurrent, content, and factorial validity (Carron & Brawley, 2000; Carron, Brawley, & Widmeyer, 1988; Carron & Spink, 1992; Estabrooks & Carron, 2000; Spink & Carron, 1992, 1993, 1994). A second condition is that the choices on an item's measurement scale must be approximately equally spaced so that all items on a scale have an equal opportunity of being chosen. The GEQ and the PAGEQ also satisfy this second condition.

James and colleagues (1984) provided formulae for nonadjusted and adjusted index-of-agreement values. Nonadjusted index-ofagreement values are computed using the average item variance for the four dimensions of cohesion—referred to as the systematic variance—in relation to the total variance. The total, or expected variance, is calculated based on the assumption that all nine options of the measurement scales have an equal likelihood of being chosen—that is, they represent what James and his colleagues referred to as a rectangular distribution. Adjusted index-of-agreement values, on the other hand, take into account the potential systematic bias that may be present in responses because of social desirability, response set, or positive leniency.

The potential level of response bias present in any given group depends on the nature of the investigation (James et al., 1984). A skewed distribution of responses, for example, is not ipso facto evidence of response bias. Raters judging an exceptional stimulus or person could reasonably be expected to use only the upper portion of a rating scale. Carron, Brawley, Eys, and colleagues (2003) noted that cohesion is a positive group property and individuals who retain their membership in a group are not likely to consistently use all nine options on a 9-point response scale; participants who perceive their group in a highly negative light are likely to remove themselves from the group. Thus, they suggested that the nonadjusted index-of-agreement values best reflects the extent of shared beliefs about cohesion because the presence of a skewed distribution does not likely reflect biases such as positive leniency but rather the state of the situation being judged. Furthermore, correcting for response bias where a skewed distribution reflects the generally positive judgments of group members could lead to an underestimation in the amount of agreement actually present. Therefore, only the nonadjusted index-of-agreement values were calculated and reported for the exercise classes in the current study.

#### RESULTS

#### **QUESTIONNAIRE TYPE**

Table 1 contains an overview of the index-of-agreement values for the four cohesion scales (i.e., ATG-T, ATG-S, GI-T, and GI-S). Prior to testing the various hypotheses, it was necessary to determine whether the level of consensus differed significantly as a function of the questionnaire used (i.e., GEQ vs. the PAGEQ). Therefore, a 2 × 4 (questionnaire type × cohesion measures) MANOVA was conducted with index-of-agreement values representing the dependent variable. The results indicated that a significant difference was present for the index-of-agreement values obtained from the two questionnaires, Wilks's lambda, F(4, 125) =63.23, p < .001;  $\eta^2 = .67$ . For all four cohesion scales, the mean index-of-agreement values obtained from the GEQ were significantly (p < .001) lower than those obtained from the PAGEQ. Thus, the total data set was subdivided into a GEQ sample (n = 31 exercise classes) and a PAGEQ sample (n = 99 exercise classes), and the various hypotheses were tested independently with each data set.

#### **OVERALL INDEX OF AGREEMENT**

*GEQ*. As Table 1 shows, the index-of-agreement values obtained from exercise classes that completed the GEQ varied from  $r_{wg(j)} = .369$  (ATG-S) to .738 (ATG-T). To compare the index of agreement across the four cohesion scales, a repeated measures ANOVA was conducted; the *F* value was significant, F(3, 28) = 14.80, p < .001;  $\eta^2 = .61$ . Post hoc Bonferroni analyses showed that the index-of-agreement values for ATG-S were significantly less (p < .05 at least) than the index-of-agreement values calculated for the other three GEQ dimensions. No significant differences (p > .05) were present among these latter three cohesion dimensions.

*PAGEQ.* As shown in Table 1 (and noted above), the index-ofagreement values were substantially and significantly larger for the exercise classes that completed the PAGEQ—varying from  $r_{wg(j)} =$ .845 (GI-S) to .962 (ATG-T). As was the case with the data from the GEQ, a repeated measures ANOVA was computed to compare the index of agreement across the four cohesion scales. The *F* value was again significant, *F*(3, 96) = 31.65, *p* < .001;  $\eta^2 = .50$ . Post hoc Bonferroni analyses indicated that the index-of-agreement values for ATG-T were significantly larger (*p* < .001) than the index-ofagreement values for the other three dimensions of the PAGEQ. The post hoc analyses also showed that the index-of-agreement values were significantly greater for GI-T than for GI-S (*p* < .001).

*Overview.* Overall, these results supported the hypothesis that the index-of-agreement values—although they could vary across the dimensions of cohesion examined—would be in a moderate range (i.e.,  $r_{wg[j]}$  from 0.40 to 0.60). With the exception of the GEQ dimensions ATG-S,  $r_{wg(j)} = .369$ , and GI-S,  $r_{wg(j)} = .572$ , the index-of-agreement values calculated for the scales of both questionnaires actually exceeded the hypothesized moderate range (i.e., were

	GEQ		PAG	PAGEQ	
Variable	М	SD	М	SD	
Individual attractions to the group-task (ATG-T)	0.738	0.24	0.962	0.04	
Individual attractions to the group-social (ATG-S)	0.369	0.30	0.882	0.18	
Group integration-task (GI-T)	0.683	0.25	0.925	0.04	
Group integration-social (GI-S)	0.572	0.25	0.845	0.14	

TABLE 1: Index-of-Agreement Values for the Scales of the Group Environment<br/>Questionnaire (GEQ; n = 31 classes) and the Physical Activity Group<br/>Environment Questionnaire (PAGEQ; n = 99 classes)

NOTE: Higher scores reveal greater consensus.

greater than  $r_{wg[j]} = .60$ ). For the most part, these findings failed to support the hypothesis that GI-T would show the strongest level of agreement and that ATG-S would show the weakest level of agreement. For both questionnaires, the ATG-T scale showed the strongest agreement; on the PAGEQ, the GI-S dimension showed the weakest level of agreement. Consistent with the hypothesis, the results indicated that the GEQ dimension ATG-S showed the weakest level of agreement in comparison to the other three dimensions of the GEQ. For both questionnaires, GI-T—which was predicted to show the greatest levels of agreement—ranked second highest in terms of consensus.

#### BETWEEN-GROUP ANALYSES

As indicated above, the index of agreement reflects the degree of consensus within each exercise class about the level of cohesion present. The presence of consensus is a necessary but not sufficient criterion for concluding that a group-level variable exists. A second important criterion is if the proportion of between-group variance is large relative to the within-group variance. To contrast the proportion of within- and between-group variance, intraclass correlation coefficients (Kenny & La Voie, 1985) and eta-squared statistics (Georgopolous, 1986) were computed. These two statistics reflect the consistency of variance in responses among members of groups nested within a larger sample in relation to nongroup members. Although theoretically, the intraclass correlation coefficient

TABLE 2:	Means, <sup>a</sup> Standard Deviations, Intraclass Correlation Coefficients (ICC),
	and Eta-Squared Values for the Four Measures of Cohesion Obtained
	With the Group Environment Questionnaire (GEQ)

Variable	GE Q			
	М	SD	ICC	$\eta^2$
Individual attractions to the group-task (ATG-T)	7.60	0.67	.17*	.22
Individual attractions to the group-social (ATG-S)	5.23	0.70	.13*	.18
Group integration-task (GI-T)	5.64	1.00	.33*	.37
Group integration-social (GI-S)	4.57	0.86	.17*	.22

NOTE: a. Cohesion scores are obtained on a 9-point scale with lower scores reflecting reduced perceptions of cohesiveness.

\*p < .05.

can range from -1.00 to +1.00, the values typically fall in the range between zero and +1.00 (Kenny & La Voie, 1985). An intraclass correlation coefficient of 1.00 would indicate that all members of the same exercise class had identical scores and these differed significantly from the scores of members of other exercise classes (with the members of each of those other classes also having identical scores). Conversely, an intraclass correlation coefficient of zero would indicate that the scores of members belonging to one exercise class were no more similar to one another than to those of members of other exercise classes. Myers (1972) proposed that a liberal criterion (i.e., p < .25) be used to test for group phenomenon rather than the typical conventional criterion (i.e., p < .05).

The eta-squared statistic provides complimentary evidence of group perceptions. Typically, values of  $\eta^2 = .20$  or greater have been used as a benchmark to support the conclusion that a group phenomenon is present (Dirks, 2000; Jehn & Shah, 1997).

*GEQ*. Table 2 provides an overview of the means, standard deviations, intraclass correlations, and eta-squared values for the four measures of cohesion obtained with the GEQ. The intraclass correlations for each of the four cohesion measures were statistically significant at p < .05. In addition, with the exception of ATG-S, the eta-squared values were all above the criterion of  $\eta^2 = .20$ . Finally, the average level of cohesion across the four scales was in the midto upper range of the 9-point scale varying from 4.57 (GI-S) to 7.60

TABLE 3:	Means," Standard Deviations, Intraclass Correlation Coefficients (ICC)
	and Eta-Squared Values for the Four Measures of Cohesion Obtained
	With the Physical Activity Group Environment Questionnaire (PAGEQ)

Variable	PAGEQ			
	М	SD	ICC	$\eta^2$
Individual attractions to the group-task (ATG-T)	7.92	0.49	.13*	.21
Individual attractions to the group-social (ATG-S)	7.09	1.02	.25*	.31
Group integration-task (GI-T)	6.94	0.81	.25*	.32
Group integration-social (GI-S)	6.13	1.10	.32*	.35

NOTE: a. Cohesion scores are obtained on a 9-point scale with lower scores reflecting reduced perceptions of cohesiveness. \*p < .05.

(ATG-T). These mean values represent a less important but nonetheless useful criterion that a group-level construct is present.

*PAGEQ.* Table 3 contains a summary of the means, standard deviations, intraclass correlations, and eta-squared values for the four measures of cohesion obtained with the PAGEQ. All of the intraclass correlations were statistically significant (p < .05), all four eta-squared statistics were above the criterion of  $\eta^2 = .20$ , and the mean values for cohesion were in the mid- to upper range of the 9-point scale.

*Overview.* The three sets of results (i.e., intraclass correlations, eta-squared values, and mean values) from analyses of both questionnaires (i.e., GEQ and PAGEQ) considered as a composite support a conclusion that exercise classes are true groups.

#### DISCUSSION

The main purpose of the current study was to examine whether perceptions of cohesion in exercise classes demonstrate sufficiently high consensus and between-group variance to support a conclusion that they are groups. Insofar as one aspect of the main purpose is concerned, there has been considerable discussion concerning what degree of consensus is sufficient to support a conclu-

sion that a collection of individuals is a group. In their discussion of this issue, Guzzo, Yost, Campbell, and Shea (1993) and Moritz and Watson (1998) suggested that values between 0.50 and 0.80 are sufficient for research purposes. In an analysis of cohesion in sport teams, Carron, Brawley, Eys, and colleagues (2003) proposed that index-of-agreement values of .40, .50, and .60 for the ATG-T and ATG-S scales of the GEQ, and values of .50, .60, and .70 for the GI-T and GI-S scales of the GEQ represent small, moderate, and large amounts of consensus. Using the Carron, Brawley, Eys, and colleagues criteria as a guide, the results from the current study showed that with the exception of one GEQ dimension (ATG-S), all index-of-agreement values calculated for the scales of the GEQ and the PAGEQ were in the moderate to high range. Thus, for the most part, the index-of-agreement values actually exceeded our a priori hypothesis that values would be in a moderate range (i.e.,  $r_{\rm wgfil}$  from 0.40 to 0.60).

Insofar as a second aspect of the main purpose is concerned, there has been general agreement on what degree of between-group differences should be present to support a conclusion that collections of individuals are groups. Analyses in the current study showed that data from the GEQ and PAGEQ easily satisfied the statistical benchmarks considered necessary in regard to the intraclass correlation coefficients (Myers, 1972) and eta-squared (Dirks, 2000; Jehn & Shah, 1997). Thus, consistent with the a priori hypothesis, the differences between exercise classes considered in concert with the level of consensus within classes offer sufficient statistical support for the conclusion that exercise classes are true groups.

Two findings pertaining to the amount of consensus in cohesion present in the two cohesion questionnaires (i.e., GEQ and PAGEQ) should be highlighted. The first is that for the two cohesion inventories, the greatest amount of consensus was shown for ATG-T—and not for GI-T. This finding was contrary to our a priori hypothesis. On the one hand, this result runs contrary to the proposition advanced by Klein and colleagues (2001) that higher consensus should be expected when group members focus on collective experiences (i.e., the *we*, *us*, and *our* cognitions queried in the GI-T items) rather than idiosyncratic experiences (i.e., the *I*, *my*, and *me* cognitions queried in the ATG-T items). On the other hand, previous research in exercise classes has consistently shown that ATG-T is the most salient of the four dimensions of cohesion (e.g., Estabrooks, 1999; Spink & Carron, 1993). Possibly, the consistently high index-of-agreement values shown for ATG-T serve to reinforce the fact that exercise classes are task oriented and participants are there for not only personal but also common reasons.

A second finding that should be highlighted is that the index-ofagreement values obtained from the PAGEQ were significantly greater than the values obtained from the exercise specific version of the GEQ. The PAGEQ was developed by Estabrooks and Carron (2000) because elderly exercisers stated that they found the negatively worded items in the exercise-specific version of the GEQ difficult to interpret. One possible explanation for the current finding (i.e., consensus was substantially higher for the exercise classes that completed the PAGEQ) may be related to the consistently positive wording of the items on the questionnaire.

With the exception of the current study, the area of shared beliefs in exercise classes remains unexplored. Thus, there are several important questions that have yet to be answered. One question relates to the issue of gender. Gender has been shown to influence a number of group dynamics factors including, for example, leadership, goal orientation, and conformity behavior (Carron & Hausenblas, 1998). Although little is known about the impact of gender on exercise class cohesion, it would be worthwhile to investigate whether the level of consensus regarding cohesion differs for all-male versus all-female exercise classes.

It should be noted that nearly 84% of the participants in the current sample were women with a mean age of approximately 55 years. Thus, the results from the current study generalize largely to people of this demographic. In the future, it would be beneficial to explore the issue of consensus using samples that consisted of younger female participants, as well as samples of older and younger male participants.

Secondly, Carron and Spink (1995) demonstrated that the size of the group is inversely related to perceptions of cohesion in exercise classes. Thus, it would beneficial to explore the impact of class size on the level of consensus present about cohesion in exercise classes. (Because our data were obtained from several different sources and we had no direct measure of the proportion of class members who completed the questionnaires, we were unable to address this issue.)

To date, research has provided support for the relationship between class leadership and various cognitive, affective, and behavioral factors including exercise adherence (Carron, Hausenblas, & Estabrooks, 2003). Therefore, another meaningful question relates to whether exercise class participants share beliefs about other important group factors such as the effectiveness of the exercise instructor and/or the quality of music in the exercise class.

Another issue of interest is whether consensus regarding cohesion changes over time. Burke (2003) addressed this issue, and the results showed that the average level of consensus about cohesion in exercise classes did not change significantly during the course of a 13-week exercise period for the four scales of the GEQ. Burke suggested that a 6-week between-test-time period (i.e., between the 4th and 10th weeks) may not have been sufficient to produce changes in the level of consensus about cohesion.

A final question relates to the value of the current research (and by extension, to the value of the other research questions just posed); that is, is consensus about cohesion in exercise classes important? One way to answer this question would be to examine whether the level of agreement about cohesion correlates with important affective states (e.g., satisfaction), cognitive states (e.g., efficacy), and behaviors (e.g., adherence).

Prior to our research, it was probable that theoreticians and practitioners would consider exercise classes to be unorganized aggregates within McGrath's (1984) typology (i.e., an aggregate of individuals who are in physical proximity while attending to a common set of stimuli). However, our results provide statistical evidence to support a conclusion that they are true groups.

#### NOTES

1. In an analysis of cohesion in sport teams, Carron, Brawley, Eys, and colleagues (2003) proposed that index-of-agreement values of 0.40, 0.50, and 0.60 for the ATG-T and ATG-S scales of the Group Environment Questionnaire (GEQ), and values of 0.50, 0.60, and 0.70 for the GI-T and GI-S scales of the GEQ represent small, moderate, and large amounts of consensus, respectively. These guidelines were used in the current study.

2. It should be reiterated that the sample consisted of 130 classes; the average age within the total sample of 130 classes was 57.40 years (*SD* = 19.92).

#### REFERENCES

- 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.
- Burke, S. M. (2003). An examination of shared beliefs in exercise classes. Unpublished master's thesis, The University of Western Ontario, London, Ontario, Canada.
- Carron, A. V., & Brawley, L. R. (2000). Cohesion: Conceptual and measurement issues. Small Group Research, 31, 89-106.
- Carron, A. V., Brawley, L. R., Eys, M. A., Bray, S., Colman, M., Dorsch, K., et al. (2003). Do individual perceptions of group cohesion reflect shared beliefs? An empirical analysis. *Small Group Research*, 34, 468-496.
- Carron, A. V., Brawley, L. R., & Widmeyer, W. N. (1998). The measurement of cohesiveness in sport groups. In J. L. Duda (Ed.), Advances in sport and exercise psychology measurement (pp.213-226). Morgantown, WV: Fitness Information Technology.
- Carron, A. V., & Hausenblas, H. A. (1998). Group dynamics in sport (2nd ed.). Morgantown, WV: Fitness Information Technology.
- Carron, A. V., Hausenblas, H. A., & Estabrooks, P. A. (2003). The psychology of physical activity. New York: McGraw-Hill.
- Carron, A. V., Hausenblas, H. A., & Mack, D. E. (1996). Social influence and exercise: A meta-analysis. *Journal of Sport and Exercise Psychology*, 18, 1-16.
- Carron, A. V., & Spink, K. S. (1992). Internal consistency of the Group Environment Questionnaire modified for an exercise setting. *Perceptual and Motor Skills*, 74, 1075-1078.
- Carron, A. V., & Spink, K. S. (1993). Team building in an exercise setting. Sport Psychologist, 7, 8-18.
- Carron, A. V., & Spink, K. S. (1995). The group size-cohesion relationship in minimal groups. Small Group Research, 26, 86-105.
- Carron, A. V., Widmeyer, W. N., & Brawley, L. R. (1985). The development of an instrument to assess cohesion in sport teams: The Group Environment Questionnaire. *Journal of Sport Psychology*, 7, 244-266.
- 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.
- Dirks, K. T. (2000). Trust in leadership and team performance: Evidence from NCAA basketball. *Journal of Applied Psychology*, 85, 1004-1012.

- Estabrooks, P. A. (1999). *The role of group cohesion in the exercise behaviour of older adults*. Doctoral dissertation, University of Western Ontario, London, Canada.
- Estabrooks, P. A., & Carron, A. V. (1999a). Group cohesion in older adult exercisers: Prediction and intervention effects. *Journal of Behavioral Medicine*, 22, 575-588.
- Estabrooks, P. A., & Carron, A. V. (1999b). The influence of the group with elderly exercisers. Small Group Research, 30, 438-452.
- Estabrooks, P. A., & Carron, A. V. (2000). The Physical Activity Group Environment Questionnaire: An instrument for the assessment of cohesion in exercise classes for older adults. *Group Dynamics*, 4, 230-243.
- Festinger, L., Schachter, S., & Back, K. (1950). Social pressures in informal groups. New York: Harper & Brothers.
- Georgopolous, D. B. (1986). Organizational structure, problem solving, and effectiveness. San Francisco: Jossey-Bass.
- 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.
- Hill, J. L., & Estabrooks, P. A. (2000, September). The relationships between group interaction and cohesion in exercise classes for older adults. Paper presented at the meeting of the Association for the Advancement of Applied Sport Psychology, Nashville, TN.
- Hill, J. L., Estabrooks, P. A., & Milliken, G. (2001). The development of class cohesion in physical activity groups for older adults. *Annals of Behavioral Medicine*, 23, S090.
- 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.
- Jehn, K. A., & Shah, P. P. (1997). Interpersonal relationships and task performance: An examination of mediation processes in friendship and acquaintance groups. *Journal of Personality and Social Psychology*, 72, 775-790.
- Kenny, D. A., & La Voie, L. (1985). Separating individual and group effects. *Journal of Personality and Social Psychology*, 48, 339-348.
- 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.
- Loughead, T. M., & Carron, A. V. (2004). The mediating role of cohesion in the leader behavior-satisfaction relationship. *Psychology of Sport and Exercise*, 5, 355-371.
- Loughead, T. M., Colman, M. M., & Carron, A. V. (2001). Investigating the mediational relationship of leadership, class cohesion, and adherence in an exercise setting. *Small Group Research*, 32, 558-575.
- Loughead, T. M., Patterson, M. M., & Carron, A. V. (2004). The impact of group factors on mood. Manuscript submitted for publication.
- Loughead, T. M., Patterson, M. M., & Carron, A. V. (2003). The relationship between cohesion and affect in physical activity groups. Manuscript in preparation.
- McGrath, J. E. (1984). *Groups: Interaction and performance*. Englewood Cliffs, NJ: Prentice Hall.

- 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.
- Myers, J. L. (1972). Fundamental in experimental design (2nd ed.). Boston, MA: Allyn & Bacon.
- Nunnally, J. C. (1978). Psychometric theory. New York: McGraw-Hill.
- Rosenkranz, S. R. (2002). *Leadership in physical activity classes for older adults*. Unpublished master's thesis, Kansas State University, Manhattan.
- 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.
- Spink, K. S., & Carron, A. V. (1992). Group cohesion and adherence in exercise classes. Journal of Sport and Exercise Psychology, 14, 78-86.
- Spink, K. S., & Carron, A. V. (1993). The effects of team building on the adherence patterns of female exercise participants. *Journal of Sport and Exercise Psychology*, 15, 39-49.
- Spink, K. S., & Carron, A. V. (1994). Group cohesion effects in exercise classes. Small Group Research, 25, 26-42.
- Widmeyer, W. N., Brawley, L. R., & Carron, A. V. (1985). Measurement of cohesion in sport teams: The Group Environment Questionnaire. London, Canada: Spodym Publishers.

Shauna M. Burke is a Ph.D. student in the School of Kinesiology at the University of Western Ontario in London, Ontario. Her research area is the psychology of sport and exercise and focuses primarily on group dynamics and physical activity.

Albert V. Carron is a professor in the School of Kinesiology at the University of Western Ontario in London, Ontario. His research focuses on group dynamics in sport teams and exercise groups.

Paul A. Estabrooks, Ph.D., is a research scientist in the Clinical Research Unit at Kaiser Permanente Colorado. His current research interests focus on utilizing group dynamics and other behavioral principles to improve clinical prevention initiatives and health care delivery. He has previously been published in Small Group Research, Journal of the American Medical Association, and the American Journal of Public Health.

Jennie L. Hill, MSc, is the research manager for Kaiser Permanente Colorado's Senior Programs. Her research interests include social and physical environmental influences on behavior, health promotion, and disease prevention. She recently won the Graduate Student Research Award from the North American Society for the Psychology of Sport and Physical activity for her work examining the development of group cohesion in exercise classes for older adults.

Todd M. Loughead is an assistant professor in the Department of Kinesiology and Physical Education at McGill University in Montreal, Quebec, Canada. He teaches undergraduate and graduate courses in the area of exercise and sport psychology. His general area of interest in research is exercise and sport psychology, with a focus

on group dynamics. Currently, his research interests are investigating the correlates of cohesion such as mood, motivation, and adherence, and the nature of peer leadership.

Michelle M. Patterson completed her Ph.D. at the University of Western Ontario. Her general area of interest is group dynamics in the sport and exercise context.

Sara R. Rosenkranz, MSc, is an administrator with USA Triathlon. Her research interests include examining group processes that occur in exercise classes. She participated in this work as part of her graduate work that examined leadership and group cohesion in exercise classes for older adults.

Kevin S. Spink is a professor in the College of Kinesiology at the University of Saskatchewan, Saskatoon. His research interests focus on group dynamics in the exercise and sport setting.