Journal of Applied Psychology
2003, Vol. 88, No. 6, 989–1004

0021-9010/03/$12.00 DOI: 10.1037/0021-9010.88.6.989

Cohesion and Performance in Groups: A Meta-Analytic Clarification of Construct Relations

Daniel J. Beal
Purdue University

Robin R. Cohen
Avon Products

Michael J. Burke and Christy L. McLendon
Tulane University

Previous meta-analytic examinations of group cohesion and performance have focused primarily on contextual factors. This study examined issues relevant to applied researchers by providing a more detailed analysis of the criterion domain. In addition, the authors reinvestigated the role of components of cohesion using more modern meta-analytic methods and in light of different types of performance criteria. The results of the authors' meta-analyses revealed stronger correlations between cohesion and performance when performance was defined as behavior (as opposed to outcome), when it was assessed with efficiency measures (as opposed to effectiveness measures), and as patterns of team workflow became more intensive. In addition, and in contrast to B. Mullen and C. Copper's (1994) meta-analysis, the 3 main components of cohesion were independently related to the various performance domains. Implications for organizations and future research on cohesion and performance are discussed.

Throughout the history of organizational research, an important goal has been to identify the factors and processes that give rise to increased group performance. In the pursuit of this goal, researchers often have focused on the social and motivational forces that exist between group members. The theoretical and intuitive hypothesis has been that these forces create a bond, or cohesion, among the members of the group, and that the stronger the bond, the greater the productivity of the group. Presumably, when cohesion is strong, the group is motivated to perform well and is better able to coordinate activities for successful performance (Cartwright, 1968; Davis, 1969). Although most researchers have acknowledged the plausibility of the relation between group cohesion and group performance, empirical observations of the relation have varied greatly, causing some authors to doubt the generalizability of the effect (Stogdill, 1972; Tziner, 1982) or to dismiss it altogether (Steiner, 1972).

Contributing to the ambiguity of the cohesion–performance relation are a wide variety of conceptualizations for both constructs. A prominent confusion, for example, concerns the appropriate level of analysis. Many researchers have measured group cohesion as individual perceptions of the group and related them to individual aspects of performance. It is not our position that such relations are insubstantial; however, because the construct of cohesion refers to the resultant of forces acting on the group, it seems appropriate to conceptualize the constructs of interest at the group level. Indeed, Gully, Devine, and Whitney (1995) tackled this very issue and found that relations between cohesion and performance were stronger when both constructs were measured at the group level. Moreover, many researchers have noted that discrepancies between conceptual and operational levels of analysis can result in ambiguous and inaccurate findings (Klein, Dansereau, & Hall, 1994; Ostroff, 1993; Scullen, 1997). Therefore, our meta-analyses include only those studies that measured cohesion and performance at the group level.

Because of the apparent ambiguity in the relation between group cohesion and performance, several meta-analyses have attempted to highlight situations in which the effect is stronger or weaker (e.g., Carron, Colman, Wheeler, & Stevens, 2002; Evans & Dion, 1989; Gully et al., 1995; Mullen & Copper, 1994; Oliver, Harman, Hoover, Hayes, & Pandhi, 1999). These studies have succeeded in identifying several moderators of cohesion–performance relationships, including group size, group reality, level of analysis, and group interdependence.

One limitation of the previous cohesion–performance meta-analyses is a lack of clarity concerning the conceptual nature of the constructs. A primary purpose of the current meta-analysis was to evaluate what we mean when we say *cohesion*, what we mean when we say *performance*, and how these construct domains relate to each other. The question of what is meant by cohesion has indeed been addressed in one of the existing meta-analyses.

Daniel J. Beal, Military Family Research Institute, Purdue University; Robin R. Cohen, Avon Products, New York; Michael J. Burke, A. B. Freeman School of Business, Tulane University; Christy L. McLendon, Department of Psychology, Tulane University.

This work was partially supported by a Robert E. Flowerree Summer Research Fellowship. We thank Howard Weiss and Janice Kelly for their comments on drafts of this article, and Jenn Kaufman, Beth Deitch, Cathy Maraist, Kim O'Farrell, and Carmen Turillo for their assistance in acquiring and categorizing studies. Portions of this article were presented at the 1st Annual Conference of the Society for Personality and Social Psychology, Nashville, Tennessee, February 2000.

Correspondence concerning this article should be addressed to Daniel J. Beal, Military Family Research Institute, Purdue University, 14 North 2nd Street, Suite 300, Lafayette, Indiana 47901. E-mail: dbeal@purdue.edu

Mullen and Copper (1994) examined the long-held notion that cohesion is composed of interpersonal attraction, group pride, and task commitment. They concluded that task commitment was significantly related to performance and that interpersonal attraction and group pride were not independently related to performance. For several reasons, however, we believe that these conclusions may have been unwarranted. Thus, one purpose of this meta-analysis is to reexamine the role of components of cohesion. In addition, we give consideration to an often neglected topic in research on cohesion and performance—the criterion domain. We have chosen to focus on these issues because a closer examination of predictor and criterion construct issues is essential for gaining a better understanding of cohesion–performance relationships. Finally, we address the particular patterns of workflow within teams and how these patterns might affect cohesion–performance relations.

Thus, the purposes of this study are to (a) conceptually reconsider the structure and content of criteria used within group cohesion studies, (b) meta-analytically test hypothesized cohesion–performance relationships with respect to more refined criterion categories, (c) constructively reexamine the independent contributions of interpersonal attraction, group pride, and task commitment in relation to criteria employed within group cohesion studies, and (d) examine the potential influence of workflow patterns on cohesion–performance relations. The remainder of the introduction unfolds as follows: First, we discuss issues concerning the structure and content of criteria in group cohesion studies. Second, we discuss psychometric and statistical reasons for reconsidering Mullen and Copper's (1994) conclusion regarding components of cohesion. Third, we describe more fully the workflow construct and explain why it might be relevant for research on cohesion and performance. Finally, we present the specific study hypotheses concerning the influence of each construct on existing cohesion–performance effects.

Cohesion and the Criterion Problem

One common feature of most of the moderators examined thus far in the cohesion–performance literature is the focus on the predictor side of the relation. This bias should come as no surprise to researchers in applied psychology; admonishments for an overemphasis on the predictors of performance, as opposed to the performance domain itself, have appeared at regular intervals throughout the history of this discipline (Austin & Villanova, 1992; Flanagan, 1956; Smith, 1976). In essence, cohesion researchers have relegated performance or criteria used within group cohesion studies to nothing more than outcomes of group cohesion. To some extent, this fractured view of the criterion domain is understandable. That is, a particularly strong argument for the benefits of group cohesion could be made if the criterion domain were inclusive of all definitions and types of criteria. Unfortunately, as we have already noted, this result has not been the case. As such, we feel that the cohesion–performance literature has a definite need for a critical examination of the criterion domain. Toward this goal, our critique of the criterion domain in the cohesion literature focuses on two main areas: whether the criterion is conceived of as a behavior or an outcome and whether output criterion measures are adjusted for inputs.

Behavior Versus Outcome

Recent treatments of the criterion domain in the applied psychological literature have noted that traditional, global conceptualizations of performance are fraught with difficulties (e.g., see Campbell, McCloy, Oppler, & Sager, 1993). Appropriate specification of criteria in applied studies rarely is given consideration beyond tacit acknowledgment as the outcome variable. In the cohesion literature, if multiple criterion measures are used, the result typically is confusion that cohesion was related to one measure and not to another. Conclusions drawn from such research frequently cite the tenuous nature of the cohesion–performance relation as the cause of the mixed results (e.g., Keyton & Springston, 1990), or the differential findings are overlooked (e.g., Cohen, Whitmyre, & Funk, 1960; Deep, Bass, & Vaughan, 1967). For these reasons, we felt that using a more fine-grained approach for the criterion domain might shed some needed light on how variables such as cohesion vary in their relation to different types of criteria.

In particular, Campbell and his colleagues (Campbell, 1990; Campbell et al., 1993) argued for a distinction between performance as behavior and performance as outcome. Put simply, performance is in the doing, not in the result of what has been done. The latter view of performance as outcome is fairly common in many areas of applied psychology, including the literature on cohesion and performance. As Campbell and others have pointed out, this latter view of performance does not take into consideration the many potential impediments to performance that are outside the control of the individual or group of individuals. For example, group sales outcomes might depend heavily on location, time of year, and economic conditions—none of which are indicators of a group's ability to perform. Because of the possibility of such impeding factors, we hypothesized that cohesion would have a stronger relation to performance behaviors than performance outcomes.

Effectiveness Versus Efficiency

In addition to differentiating between behaviors and outcomes, we also wanted to examine how the consideration of inputs affected cohesion–performance relationships. When comparing groups, efficiency measures, which adjust for group inputs, often are more informative of performance than effectiveness measures, which only permit a comparison of group outputs (e.g., see Borucki & Burke, 1999; Wilderom, Glunk, & Maslowski, 2000). For example, if a retail organization with stores of many different sizes wanted to evaluate the performance of each store, overall effectiveness measures (e.g., gross sales) would not provide an accurate assessment of performance. A larger store most likely would rate higher simply because of its greater capacity to produce. A smaller store, despite extremely efficient use of its resources, might never be able to reach the overall level of performance of the larger stores. Efficiency measures, which take inputs as well as outputs into account, might better reflect the true nature of store performance.

Aside from the influence of measurement issues of effectiveness and efficiency, there are theoretical reasons to believe that cohesion might bear a stronger relation to efficiency measures than to effectiveness measures. A variety of group researchers have posited that cohesion is an important variable linking group processes and group outcomes. Although the temporal placement of cohesion in this causal process is uncertain, researchers have found cohesive groups to have increased efficiency of language behavior (Mickelson & Campbell, 1975), greater team mental model convergence (Mathieu, Heffner, Goodwin, Salas, & Cannon-Bowers, 2000), and greater use of transactive memory systems (Hollingshead, 1998, 2000; Wegner, Erber, & Raymond, 1991). In sum, cohesive groups should be able to use their groups' resources more efficiently because they know the members of the group better and are motivated to complete the task successfully. Effectiveness measures may be somewhat insensitive to this aspect of performance, but efficiency measures are designed to hone in on these sorts of processes. As such, we hypothesized that relations between cohesion and efficiency measures would be greater than relations between cohesion and effectiveness measures.

The particular reasons for expecting this pattern of cohesion–performance relations highlights another major purpose of this meta-analytic integration. Our arguments supporting moderating roles for behavior versus outcome and efficiency versus effectiveness are based on the assumption that interpersonal attraction, task commitment, and group pride are all important aspects of group cohesion. This view is at odds with a previous meta-analysis of the components of cohesion (Mullen & Copper, 1994). These authors examined several components of cohesion and concluded that task commitment (as opposed to interpersonal attraction and group pride) was the only component that independently contributed to the cohesion–performance relation. This finding was surprising, especially considering that most of the research on cohesion and performance has operationalized cohesion almost completely in terms of interpersonal attraction (Lott & Lott, 1965). After briefly reviewing some previous perspectives on the dimensionality of cohesiveness, we present three arguments suggesting that Mullen and Copper's (1994) conclusions may have been premature.

Components of Cohesion and the Need to Reconsider Mullen and Copper's (1994) Conclusions

The recognition of cohesion as a multidimensional construct dates back to the seminal work of Festinger (1950). He discussed cohesion as a culmination of factors, such as attraction to the members of a group, the activities of a group, and the prestige of the group. Despite this early emphasis on acknowledging the components of cohesion, researchers often have measured only one aspect of the construct (e.g., Seashore, 1954) or have used an omnibus measure that cannot determine the independent contributions of each component (e.g. Gowda, 1988). This fact is not to say, however, that the importance of components of cohesion has been completely ignored. In response to several articles lamenting the unidimensional approach to measuring cohesion (Carron, 1982; Mudrack, 1989; Tziner, 1982), numerous researchers have explored relations between its components and performance (Carron, Widmeyer, & Brawley, 1985; Zacarro & Lowe, 1986; Zacarro & McCoy, 1988).

Considering the original emphasis of components in the cohesion construct and the continued emphasis on this approach (Carless & De Paola, 2000; Cota, Evans, Dion, Kilik, & Longman, 1995), Mullen and Copper's (1994, p. 224) conclusion, "that efforts to enhance group performance by fostering interpersonal attraction or 'pumping up' group pride are not likely to be effective," was somewhat surprising. Recent advances in meta-analytic methods, however, suggest several reasons why this conclusion needs to be reconsidered. Below, we discuss how issues related to levels of analysis, stochastically dependent effects, and the use of regression weights for determining the relative contributions of components of cohesion each contribute to uncertainty regarding Mullen and Copper's conclusion.

Mixed Levels of Analysis

One interpretational difficulty concerning the results of Mullen and Copper's (1994) meta-analysis is the inclusion of studies in the same distribution of effect sizes that measured the group cohesion and performance variables at an individual level, as well as studies that measured them at a group level. Gully et al. (1995) documented how this practice can lead to an underestimation of the effect because the cohesion–performance relation is weaker at the individual level. Because individual-level studies were included in the component of cohesion moderator tests, the results are ambiguous to some extent.

The problem of mixing levels of analysis is compounded if the effects are sample-size weighted, as is customary for distributional meta-analyses (Hunter & Schmidt, 1990). The primary issue in such meta-analyses is that effects measured at the individual level necessarily will be weighted more heavily than effects measured at a group level of analysis. For example, let us say that a study assessing the cohesion–performance relation was conducted, and measurements were taken both at the individual and group level of analysis. This hypothetical study used 25 groups with 5 people in each group. If both the individual-level and group-level correlations were later included in a meta-analysis, the effect obtained at the individual level of analysis, with an $N = 125$, would be weighted five times greater than the effect obtained at the group level of analysis, with an $N = 25$. Thus, not only do individual-level assessments of group cohesion and performance suffer from conceptual ambiguity, the ambiguities are amplified when correcting for sample size.

Stochastically Dependent Effects

Beyond the problems of combining different levels of analysis, there are methodological and statistical issues that add uncertainty to Mullen and Copper's (1994) analysis of components of cohesion. Advances in the understanding of the methods of meta-analysis have revealed that the use of stochastically dependent effect sizes can lead to large errors in parameter estimation (Gleser & Olkin, 1994). *Stochastically dependent* describes a situation in meta-analysis in which several estimates of an effect from the same sample are treated as separate, independent estimates. This inclusion of nonindependent effects often occurs when there are too few independent effects to conduct analyses with adequate power. For example, in Mullen and Copper's meta-analysis of components of cohesion, reasonable estimates of cohesion–performance relationships for each component of cohesion would have been difficult given the small number of independent effects for the respective components. Therefore, Mullen and Cooper took advantage of the fact that several studies had multiple estimates of the cohesion–performance relation, thus rendering an overall more stable effect size estimate. This method of analysis, however, unduly weighted the importance of studies with multiple estimates. Multiple measurements of the same cohesion–performance relation may make a more reliable estimate, but this stability is artificial because it does not take into account the nonindependence of the data. We sought to include only stochastically independent effects within each of our effect size distributions in order to obtain a more accurate estimate of relations between components of cohesion and performance.

Inappropriate Regression Weighting

A final issue with Mullen and Copper's (1994) analysis concerns the use of regression weights to examine the independent contribution of each component of cohesion in predicting performance. These regression weights were based on the relative proportion of items measuring a particular component of cohesion in a questionnaire.[1] Unfortunately, the number or proportion of items measuring each component in a questionnaire cannot provide information concerning the independence of each component's relation to performance. For example, a hypothetical cohesion–performance study examines 10 groups, each having rated cohesiveness with a four-item scale. Three of the items assess task commitment, and one item assesses interpersonal attraction. Included also is a single-item criterion measure of group performance. In this hypothetical study, the correlation between the interpersonal attraction item and performance was .80, and the correlation between the average task commitment score (i.e., the average of all three items) and performance was .20. The correlation between the overall average cohesion score (i.e., all four cohesion items) and performance was .35—an effect that is larger than most in the cohesion–performance literature. Clearly, the effect size of the overall scale is due to the one interpersonal attraction item. In Mullen and Copper's analysis, however, this effect would have been attributed mostly to task commitment because the overall scale has three times as many task commitment items. This possibility clouds Mullen and Copper's interpretation that the task commitment component of cohesion is the only independent predictor of performance.

In the current study, we circumvented this problem by including only independent estimates of each component's effect in any particular meta-analytic distribution of effects. That is, if an effect size estimate included items assessing more than one component simultaneously, we did not include it in the analysis. This strategy may not have been possible at the time of Mullen and Copper's (1994) meta-analysis because there were not sufficient numbers of studies to examine the effect size estimates separately. However, the substantial increase in the recent cohesion–performance literature since Mullen and Copper's work afforded us the luxury of such an analysis.

Patterns of Team Workflow

Throughout the history of organizational research on groups, one sentiment has been pervasive: task type matters. Whether researchers are concerned with leadership style (Weed, Mitchell, & Moffitt, 1976), group member status (Kirchler & Davis, 1986), group structure (Stewart & Barrick, 2000), or group cooperation (Kabanoff & O'Brien, 1979), task type has attained an important role. As a result, many taxonomies of task type have been proposed, emphasizing a variety of specific characteristics (e.g., Hackman, 1968; Hackman & Morris, 1975; McGrath, 1984; Steiner, 1972). Although these taxonomies have been useful for many areas of group research, few have exhibited any relevance for group cohesion. It appears, however, that the interdependence of the task may be important. In particular, Gully et al. (1995) found task interdependence to moderate cohesion–performance relations at the group level.

In the current meta-analysis, we followed up on the finding by Gully et al. (1995) by examining how specific aspects of task interdependence might interact with group cohesion and performance. Specifically, we examined how the pattern of a team's workflow can enhance the beneficial aspects of group cohesion. Thompson (1967) initially discussed the notion of internal interdependence at a more macro-organizational level. In this discussion, he defined four forms of interdependence describing how different branches of an organization exchange information and work. Tesluk, Mathieu, Zaccaro, and Marks (1997) provided a similar taxonomy that described how work flows between members of a team. Because our interest was at the team or group level, we adopted the specifics of Tesluk et al.'s taxonomy.

Conceptually, differences in the exchange of work between members of a team can vary in several ways, including the direction of workflow and the number of exchanges. For example, *pooled* workflow involves tasks that aggregate individual performances to the group level. No interactions or exchanges between group members are required in this pattern of teamwork. Work does not flow through multiple members of the group, and performance simply is the sum or some other aggregation of the group members' performances. *Sequential* workflow describes tasks that move from one member of the team to another but not in a back-and-forth manner. Group performance is not simply the pooling of each member's performance but is a function of how the work progresses through each member of the group. For example, in an assembly line, each member of a group would be responsible for a particular portion of the final product. Line assembly, however, implies that after one portion of the product is assembled, it is passed along to the next person until it reaches the end of the line. After the last person has completed his or her part, the product is complete. Thus, work flows sequentially from the first team member to the last team member.

The final two patterns of team workflow involve considerably more workflow between team members. *Reciprocal* workflow is similar to sequential in that work flows only from one member to another, but the flow is now bidirectional; team members can exchange work with one another multiple times. The team performance, however, is accomplished when the last person in the group has completed his or her performance. Finally, *intensive*

[1] For a few studies, weights were assigned on the basis of the extent to which a particular component was manipulated in experimental conditions. Our argument, however, still holds for this procedure. That is, a greater emphasis on one component in a manipulation of cohesion does not necessarily mean it held the responsibility for the relation between overall cohesion and performance.

patterns of workflow occur when the work has the opportunity to flow between all members of the group, and the entire group must collaborate to accomplish the task.

As can be seen, both the directional changes and the amount of workflow between team members increases as the patterns progress from pooled to sequential to reciprocal to intensive. To the extent that group members exchange greater amounts of work between members, group processes such as cohesion should gain an important function in contributing to performance. In discussing this notion, Tesluk et al. (1997) recommended a variety of human resource programs to improve productivity for each pattern of team workflow. As the amount of between-member workflow increased, these authors recommended a greater emphasis on team-level (as opposed to individual-level) programs. Consistent with these observations, we expected cohesion to bear stronger relations to performance as the level of team workflow increased from pooled through intensive. That is, factors such as attraction to group members, a shared commitment to the task, and a sense of pride in belonging to the group should have a greater impact on performance as the workflow between members increases in each pattern of teamwork.

Summary

Our goals for the current meta-analytic integration of cohesion–performance relations were (a) to gain a more complete understanding of cohesion–performance relations with respect to different types of criteria employed in group cohesion studies, (b) to constructively reexamine the independent contributions of interpersonal attraction, group pride, and task commitment in relation to performance, and (c) to identify what particular patterns of workflow would have benefits or detriments for the link between cohesion and performance. Our specific hypotheses regarding performance criteria were that cohesion–performance relations would be stronger when performance is conceptualized and measured as a behavior as opposed to an outcome, and measures of performance efficiency would have a stronger relation to cohesion than measures of performance effectiveness. We further hypothesized that all three components of cohesion would bear significant independent relations to performance criteria. Finally, we hypothesized that patterns of team workflow would moderate the cohesion–performance relation such that greater amounts of workflow would be associated with stronger correlations.

Method

Literature Search

Several different approaches were used in locating relevant articles. First, attempts were made to retrieve as many articles as possible from the reference sections of Mullen and Copper (1994) and Gully et al. (1995). Second, we conducted computer searches of the PsycINFO and Sociofile databases using the search terms *cohesion*, *cohesiveness*, *interpersonal attraction*, *group attraction*, *task commitment*, *task attraction*, or *group pride*. This search then was combined with a search for *productivity*, *performance*, *effectiveness*, or *efficiency*. In addition, reference sections of the final articles were scanned for any citation that might contain an estimate of the cohesion–performance relation. Finally, requests for unpublished studies were made of several experts in the field of group research, as well as the electronic mailing list for the Society for Personality and Social Psychology. This overall search returned several hundred possible articles, which then were narrowed down to 145 after discarding clearly irrelevant or unobtainable studies.

The articles then were examined individually for inclusion in the final analyses. Fifty studies were removed because (a) only multivariate or partialed effect size estimates were available, (b) the performance variable was a self- or within-group rating (see Gully et al., 1995), or (c) the operationalization of cohesion or performance did not conform to our definitions (see below). This process left 64 separate articles with 71 independent estimates and 186 total estimates of cohesion–performance relations. Of these 64 separate articles, 11 were unpublished articles or dissertations. Publication dates ranged from 1951 to 2002, with a median year of publication of 1989 (SD = 14.03). Sample sizes for the studies averaged 45.30 (SD = 33.93), and group sizes averaged 6.22 (SD = 3.66). Articles used in the meta-analysis are listed along with their uncorrected correlations in Table 1.

In comparison to the two largest of the previous broad meta-analyses, Gully et al. (1995) obtained 51 independent effect size estimates and Mullen and Copper (1994) obtained 52 independent effect size estimates. Note, however, that both of these meta-analyses included group- and individual-level effects. When considering only the group-level effects, Gully et al. contained only 35 independent effects and Mullen and Copper only 39 independent effects. Of these two meta-analyses, we included 29 of Mullen and Copper's 39 effects (74%) and 25 of Gully et al.'s 35 effects (71%). There were several reasons why the overlap was not 100%: Some effects were multivariate (i.e., included variables other than cohesion in the same analysis), some were self-reports of performance, one study was unobtainable, and the dubious quality of one study removed it from consideration in our analysis. Given these considerations, the overlap with previous meta-analytic samples seemed high. In addition, the current meta-analysis contained approximately 85% more group-level studies than the largest previous effort, which is a substantial increase.

Coding of Study Characteristics

Perhaps one of the largest sources of error in meta-analyses is the multitude of judgment calls made at various stages of the research synthesis process (Wanous, Sullivan, & Malinak, 1989). For example, Mullen and Copper's (1994) definition and subsequent coding of group interaction led them to conclude that the level of interactivity does not moderate the cohesion–performance relation. Gully et al. (1995), however, defined interaction differently (actually termed *interdependence*) and found that highly interactive groups exhibited a stronger cohesion–performance relation at the group-level of analysis. It is likely that the use of a different operational definition of group interaction is at least partly responsible for the discrepant conclusions. An important early step, therefore, was to arrive at thorough operational definitions for every relevant variable (Cooper, 1998). Provided here is a list of each variable along with how it was operationally defined. In addition, we have included examples of each variable in Appendix A.

Performance as Behavior

In this category we included effects whose criterion measure was an evaluation of actions or behaviors relevant to the goals of the study as indicated by the experimenter (i.e., in laboratory experiments) or the goals of the organization (i.e., in field studies). Following Campbell et al. (1993), this definition also included indices of unobservable cognitive behaviors

Table 1
Sample Sizes and Observed (Uncorrected) Correlations From Each Study and for Each Variable of Interest

Study	*N*	IA	TC	GP	Beh.	Outcm.	Effct.	Effic.	P	S	R	I
Arroyo, 1997	12		.349			.349	.349		.349			
Bakeman & Helmreich, 1975	10	.430			.430		.430		.430			
Barrick et al., 1998	51							.270				
Bhatara, 1972												
Study 1	40				.354			.354		.316	.392	
Study 2	20				.478			.478			.478	
Bird, 1977	8					.745	.745					.745
Blades, 1986	45					.110						
Burchfield, 1997	72	.105				.105		.105			.105	
Carpenter & Radhakrishnan, 2002	30	.174				.174	.174					.174
Carron & Ball, 1977	12	.031	.479	.116		.303	.303					.303
Cohen et al., 1960	16	.131			.131			.131	.131			
Colarelli & Boos, 1992	86	.050				.050	.050					.050
Cotter, 1979	13	.048		.042		.052	.052					.052
Craig & Kelly, 1999	61	−.095	.159			.032	.032					
Darley et al., 1952	13	.331	.277			.307	.307					.307
Deep et al., 1967	9	−.434				−.368	−.381	−.355				−.368
Duffy & Shaw, 2000	137					.200	.200					
Eisenberg, 2001	48				.268			.268	.248			
Elias et al., 1989	36		.417		.417			.417			.417	
Fiedler, 1954	22	.335				.335	.150	.520			.335	
Fodor & Smith, 1982	40		−.077		−.077		−.077		−.077			
Fox, 1986	94					.080	.080					
Gekoski, 1952	21	.100	.100			.100		.100	.100			
George & Bettenhausen, 1990	33	.040				.040		.040	.040			
Gonzalez et al. (in press)	71	.090	.260			.175	.175				.175	
Goodacre, 1951	12	.725			.725							.725
Greene, 1989	54	.150				.150						
Hemphill & Sechrest, 1952[a]	85	.178			.151	.278	.162	.360				
Hoegl & Gemuenden, 2001	145						.170	.215				
Jaffe & Nebenzahl, 1990	20	−.200	.389			.094		.094				.094
Jehn, 1994	88	.170				.170	.170					.170
Jehn & Shah, 1997	53	.555			.555		.460	.650				.555
Karau & Hart, 1998	30	.488			.488			.488	.488			
Keller, 1986	30	.435					.510	.360				
Keyton & Springston, 1990	35				.079	−.112	−.017					−.017
Klein & Mulvey, 1995												
Study 1	52	.230				.230	.230					.230
Study 2	89	.150			.150			.150		.150		
Landers et al., 1982	10	.780	.610	.540		.640	.640					.640
Langfred, 1998	61	.410	.170					.290				
Langfred, 2000												
Sample 1	67	.280				.280	.280					
Sample 2	61	−.650				−.650	−.650					
Lodahl & Porter, 1961	55	.190				.190		.190				
Lorenz, 1985	21							.469				.469
Martens & Peterson, 1971	144	.064		.158		.130	.130					.130
Melnick & Chemers, 1974	21	.220		.130		.063	.063					.063
Miesing & Preble, 1985	6					.867		.867				.867
Mossholder & Bedeian, 1983	18						.120		.120			
Mulvey & Klein, 1998												
Study 1	59	.370				.370	.370					.370
Study 2	101	.350				.350	.350					.350
Neal, 1997	25					.520	.520					.520
Neubert, 1999	21				.490			.490				
Norris & Niebuhr, 1980	18	.440				.440		.440				.440
Podsakoff et al., 1997												
Study 1	40					.260		.260				
Study 2	71					.025	.100		.025			
Porter & Lilly, 1996	80		.190			.190	.190					.190
Seers et al., 1995	6					.700						
Smith et al., 1994	53	.350				.350		.350				.350
Stinson & Hellebrandt, 1972												
Sample 1	11					.110	.110					.110
Sample 2	14					.000	.000					.000
Tehan, 1983	16					.465	.465					.465

Table 1 (*continued*)

Study	*N*	IA	TC	GP	Beh.	Outcm.	Effct.	Effic.	P	S	R	I
Terborg et al., 1976	42	−.300				−.300	−.300		−.300			
Tesluk & Mathieu, 1999	88	.265			.229	.043						
Tziner & Vardi, 1983	115	.320										.320
Wech et al., 1998	71							.200				
Wekselberg et al., 1997[a]	10				.476		.476					.476
Williams & Hacker, 1982	9	.605	.720	.770		.747	.747					.747
Wolfe & Box, 1988	36	.045				.045		.045				.045
Wong, 1992	40	.268			.268			.268		.268		
Zaccaro & Lowe, 1986	54	−.040	.439		.200			.200	.200			
Zaccaro & McCoy, 1988	33	.161	.105			.133	.133					.133
Zaccaro et al., 1995	46		.301		.301			.301			.301	

Note. IA = interpersonal attraction; TC = task commitment; GP = group pride; Beh. = performance as behavior; Outcm. = performance as outcome; Effct. = effectiveness criteria; Effic. = efficiency criteria; P = pooled workflow; S = sequential workflow; R = reciprocal workflow; I = intensive workflow.
[a] Different effects within the study had different *N*s. The overall mean *N* is provided.

(i.e., " 'solutions,' 'statements,' or 'answers' produced as a result of covert cognitive behavior and totally under the control of the individual" p. 40).[2]

Performance as Outcome

Criteria were categorized as performance outcomes if they represented the consequences or results of performance behaviors.

Measures of Effectiveness

Performance effectiveness was defined as an evaluation of the results of performance with no consideration of the costs of achieving the results.

Measures of Efficiency

Performance efficiency was defined as the effectiveness of a group with some consideration of the cost of achieving that level of effectiveness, that is, a ratio or factoring in of inputs relative to outputs. If a measure took inputs into account in any way, then it was considered an efficiency measure. We interpreted inputs in a broad manner, including time, effort, and other resources expended, as well as number of errors made and relative size of the group (if size offered performance benefits). Because time is always involved in performance, we considered time as an input if it was explicitly mentioned as part of the task (e.g., participants had 15 min to come up with solutions, workers were stopped after 12 min, and so forth).

Team Workflow

Our definitions for teamwork patterns follow directly from Tesluk et al. (1997, p. 201).

Pooled. Work and activities are performed separately by all team members, and work does not flow between members of the team.

Sequential. Work and activities flow from one member to another in the team but mostly in one direction.

Reciprocal. Work and activities flow between team members in a back-and-forth manner, but only a single team member is worked with at a given moment in time.

Intensive. Work and activities come into the team, and members must collaborate as a team in order to accomplish the team's task.

Cohesion

Variables were considered to measure cohesion if they fell into one of the following component categories:

Interpersonal attraction. A shared liking for or attachment to the members of the group.

Task commitment. The extent to which the task allows the group to attain important goals or the extent to which a shared commitment to the group's task exists.

Group pride. The extent to which group members exhibit liking for the status or the ideologies that the group supports or represents, or the shared importance of being a member of the group.

Beyond the categorizations described above, there were several other judgment calls that needed to be made. Frequently, there were multiple estimates of the cohesion–performance relation, even within a particular category (e.g., three reported correlations between interpersonal attraction and effectiveness for one particular task). In general, one overall estimate was obtained with a sample-weighted average. However, if the correlations were computed at different time periods in the group's development, the correlation between the latest occurring measurements of cohesion and performance was assumed to be the best estimate. Our reasoning for using the most recent cohesion–performance correlation concerns the nature of the construct of cohesion. Cohesion is not something that occurs immediately on a group's formation. It develops after the group has had an opportunity to work together or at least become acquainted with each other (Gosenpud, 1989; Harrison, Price, & Bell, 1998; Matheson, Mathes, & Murray, 1996). Therefore, we felt that the effects of cohesion on group performance would most likely emerge later in the group's existence as compared with earlier.

In addition, there was a question of whether a questionnaire mostly measured one particular component over others. Because we wished to examine the independent effects of different components of cohesion on performance, some effects could not be coded for the particular type of cohesion. For example, if a questionnaire mostly contained items measuring interpersonal attraction but also had one or two items measuring task commitment, the effect was not coded for component of cohesion. This

[2] Group decision-making tasks represent a source of possible confusion with our definitions. In several cases, the group decision-making process was unobservable to the person who judged the performance (i.e., only the end result was judged, not the process). Despite being "unobservable" we coded these situations as outcomes. Our reasoning was that many observable behaviors occurred in the context of the performance and the unobservable element had little to do with cognitive processes that could not be measured. Therefore, this type of performance could not fall under the "unobservable cognitive behaviors" portion of our performance behavior definition. If, however, aspects of the decision-making process were included as measures of performance, we coded these elements as behaviors.

ensured that the effects of the other components did not contaminate assessments of a particular component of cohesion.[3]

Data Analysis Procedure

Tests of the hypotheses were conducted using Raju, Burke, Normand, and Langlois' (RBNL, 1991; Finkelstein, Burke, & Raju, 1995) meta-analytic procedures, with a random effects model (cf. Raju & Drasgow, 2003). The RBNL procedure uses sample statistics including available information on sample-based artifacts (i.e., predictor and criterion reliabilities) to estimate individually corrected effects with standard errors for these corrected effects. Subsequently, this meta-analytic procedure computes sample-size weighted estimates of the mean and variance of corrected effects. In contrast, most other meta-analytic procedures (cf. Hunter & Schmidt, 1990) rely on distributions of hypothetical artifact values (as discussed in more detail in Paese & Switzer, 1988, and Raju, Pappas, & Williams, 1989) for estimating the mean and variance of corrected effects. It is noteworthy that the RBNL meta-analytic procedure permits the construction of a confidence interval around the estimated mean corrected effect (cf. Finkelstein et al., 1995). To use the RBNL procedure, all study effect sizes were first converted to Pearson correlations. These correlations were then corrected for sampling error, as well as unreliability, in the predictor and criterion measures. In instances in which a study did not report reliability information, the average reliability from available studies was substituted for the missing values.

Our first step was to constructively replicate the meta-analysis by Mullen and Copper (1994). This was done primarily to ensure that potentially differing results concerning components of cohesion were not due solely to differences in the particular studies included. To accomplish this, we analyzed separately the set of studies that overlapped with Mullen and Copper. After obtaining high levels of agreement, we proceeded to code the remaining studies (see Appendix A for indices of interrater agreement). The first author coded the variables for each study, and the fourth author cross-coded all variables for each study. Where there were disagreements, the two coders reached agreement through discussion. This double coding procedure and the checks on interrater agreement for the coding of key study characteristics are consistent with recent recommendations for coding studies in meta-analyses (see Burke & Landis, 2003). We note here that the pattern of results obtained for the Mullen and Copper replication did not differ substantially from the results of the final set of studies. Therefore, the reported results reflect our complete set of studies.

Results

Overview of Analyses

To examine our hypotheses, we focused on three main criteria: mean corrected correlation ($M_{\hat{\rho}}$), estimated *SE* of $M_{\hat{\rho}}$, and confidence intervals around $M_{\hat{\rho}}$. We chose to focus on corrected coefficients because we were primarily interested in estimated relations between well-developed, highly reliable measures of cohesion and performance. Note that we interpreted our findings using confidence intervals as opposed to credibility intervals. Confidence intervals reflect the accuracy of the mean coefficient through the use of the standard error of the corrected mean correlation (Finkelstein et al., 1995), whereas credibility intervals estimate the generalizability of the effect through the use of the standard deviation of the observed or corrected correlations (Schmidt & Hunter, 1977; Whitener, 1990).

We have not reported tests of generalizability for several reasons. First, we were not concerned with the "portability" of the effect. Previous meta-analyses have demonstrated significant heterogeneity in the cohesion–performance relation (Gully et al., 1995; Mullen & Copper, 1994). As discussed earlier, several moderating variables already have been identified, mostly concerning the predictor side of the relation. Our goal was to examine constructs that produce different mean correlations based on theoretical differences. Therefore, we examined corrected mean coefficients and the extent to which their confidence intervals overlap. A second reason why generalizability tests, such as the credibility interval, were not used is because of their notoriously inaccurate Type I error rates (Cornwell, 1993; Koslowsky & Sagie, 1993; Sackett, Harris, & Orr, 1986; Spector & Levine, 1987).

We first present results concerning different types of criteria including behavior versus outcome and effectiveness versus efficiency. Following this, we address the independent contributions of interpersonal attraction, group pride, and task commitment in predicting performance. We then examine, where possible, the interactions between components of cohesion and the various types of criteria. Finally, we address the moderating effects of team workflow.

Different Types of Criteria

Table 2 presents summary information for each construct of interest. Table 3 presents average reliabilities for each predictor and criterion category where available. Below we consider the results for each of these variables in more depth.

Behavior Versus Outcome

We considered performance behaviors to be more closely linked to the process of cohesion than performance outcomes, which often are determined by factors unrelated to the efforts of the group. Therefore, we predicted that cohesion–performance relations, on average, would be stronger when performance was operationalized as behavior than when performance was operationalized as an outcome. In support of this hypothesis, the mean corrected correlation was greater for performance behaviors ($M_{\hat{\rho}} =$ .301) than for performance outcomes ($M_{\hat{\rho}} =$ .168). Note that the confidence interval around the cohesion–behavior correlation does not include the cohesion–outcome correlation and vice versa. Indeed, these correlations were significantly different from each other ($Z = 2.573$, $p < .05$). Also of note in this analysis is that the confidence intervals of these correlations did not include zero. Thus, cohesion is related to both conceptualizations of performance but accounts for more variability in performance behavior.

[3] Although our meta-analytic estimates have eliminated stochastically dependent effects within a particular analysis, a small amount of sample dependency exists when comparing meta-analytic correlations with each other. Specifically, many studies included several effect size estimates that could be included in more than one of our moderator categories (e.g., if a study included one questionnaire that was interpersonal attraction and another that was task commitment). Thus, when comparing these two average correlations, the same effect size does not appear in both estimates, but a small percentage of the same individuals may have contributed to both estimates. To examine this issue, we reanalyzed our data after removing all studies with overlapping samples. Because the results of these analyses did not differ substantially from the original estimates, and because there were no changes in the pattern or significance of our results, we reported the original estimates.

Table 2
Mean Uncorrected Correlations ($\bar{r}$), Mean Corrected Correlations ($M_{\hat{\rho}}$), Variances of $\hat{\rho}$ ($\sigma^2_{\hat{\rho}}$), Standard Errors of the Mean of $\hat{\rho}$ ($SE_{M_{\hat{\rho}}}$), 95% Confidence Intervals (CI), Number of Effect Sizes (K), and Number of Groups (N) for Criterion and Component Moderators

Moderator	$\bar{r}$	$M_{\hat{\rho}}$	$\sigma^2_{\hat{\rho}}$	$SE_{M_{\hat{\rho}}}$	95% CI	K	N
Performance							
Behavior	.267	.301	.008	.041	.220, .383	19	778
Outcome	.147	.168	.037	.036	.096, .239	47	2,125
Measure							
Effectiveness	.155	.175	.045	.041	.095, .256	40	1,899
Efficiency	.272	.310	.006	.031	.249, .370	31	1,337
Component of cohesion							
Interpersonal attraction (IA)	.171	.199	.052	.042	.117, .281	43	2,049
Behavior	.279	.315	.028	.069	.179, .451	10	482
Outcome	.124	.139	.053	.050	.041, .237	31	1,446
Effectiveness	.132	.148	.074	.062	.026, .270	25	1,187
Efficiency	.240	.284	.032	.055	.176, .393	19	792
Task commitment (TC)	.246	.278	.000	.043	.194, .361	16	579
Behavior	.281	.302	.026	.074	.156, .448	4	176
Outcome	.242	.273	.000	.055	.166, .381	11	342
Effectiveness	.205	.232	.002	.055	.124, .340	10	341
Efficiency	.306	.343	.000	.068	.210, .475	6	238
Group pride	.242	.261	.000	.065	.133, .389	6	209

Effectiveness Versus Efficiency

We also posited that measures of performance efficiency would be particularly adept at capturing the process benefits of group cohesion. In contrast, measures of performance effectiveness, which only take outputs into account, should be less able to exhibit the beneficial effects of a cohesive group. In line with this hypothesis, efficiency measures ($M_{\hat{\rho}} = .310$) possessed a stronger average correlation than did effectiveness measures ($M_{\hat{\rho}} = .175$). These mean coefficients fell outside of the opposing correlation's confidence intervals and were significantly different from each other ($Z = 2.787$, $p < .05$). Much like the results for behavior versus outcome, note that both confidence intervals were above zero; efficiency measures better reflected the benefits of cohesion, but cohesive groups also were more effective.

Table 3
Sample-Weighted Average Reliabilities for Each Predictor and Criterion Category ($\bar{r}_{xx}$ and $\bar{r}_{yy}$)

Predictor	$\bar{r}_{xx}$	K	N
Interpersonal attraction	.845	24	1,422
Task commitment	.866	5	245
Group pride	—	—	—
Criterion	$\bar{r}_{yy}$	K	N
Performance behavior	.874	16	719
Performance outcome	.888	34	1,525
Effectiveness	.911	27	1,283
Efficiency	.894	26	1,218

Note. Dashes are indicative of the fact that no reliability coefficients were available for studies examining group pride; therefore, the analysis used the average reliability for all cohesion measures. K = number of effects reporting reliability; N = total number of groups for each average reliability.

Components of Cohesion

Most importantly for the analysis of components of cohesion, all three mean correlations were significantly greater than zero. There were differences in the magnitude between the mean effects for each component of cohesion, but none of the mean corrected effects were significantly different from each other (all Zs < 1.30, ps $> .18$). The ascending order of effect sizes was interpersonal attraction ($M_{\hat{\rho}} = .199$), group pride ($M_{\hat{\rho}} = .261$), and task commitment ($M_{\hat{\rho}} = .278$). Thus, in contrast to the findings of Mullen and Copper (1994), our analysis indicates that the three components of cohesion each correlate meaningfully with performance criteria. Also of interest is that group pride, despite being the most frequently ignored component of cohesion and having only six effect size estimates, had an average correlation that was significantly greater than zero and as strong as the other components.

Interactions Between Components of Cohesion and Types of Criteria

Although there were too few studies to break down each component of cohesion with each category of criteria, several analyses were possible. Interpersonal attraction had, by far, the largest number of studies, and we were able to examine how this component related to behavior versus outcome, as well as to effectiveness versus efficiency. Task commitment, although having substantially fewer effects than interpersonal attraction, still allowed for tests of behavior versus outcome and effectiveness versus efficiency. The handful of group pride effects, however, all used criteria of effectiveness (i.e., no measures of efficiency) and outcome (i.e., no measures of performance behavior).

Interpersonal attraction displayed a pattern similar to the overall findings for behavior versus outcome. Specifically, the interpersonal attraction–outcome relation ($M_{\hat{\rho}} = .139$) was significantly smaller than the interpersonal attraction–behavior relation ($M_{\hat{\rho}} =$

.315; $Z = 2.173$, $p < .05$). Task commitment, in contrast, did not exhibit a clear difference between behavior and outcome (behavior, $M_{\hat{\rho}} = .302$; outcome, $M_{\hat{\rho}} = .273$; $Z = .313$, $p = .754$).

Interpersonal attraction displayed the same pattern for effectiveness versus efficiency as observed in the overall analyses, although the difference now reached only marginal levels of significance (effectiveness, $M_{\hat{\rho}} = .148$; efficiency, $M_{\hat{\rho}} = .284$; $Z = 1.722$, $p = .085$). Although the same pattern was noticeable for effectiveness versus efficiency in the task commitment average correlations, the difference did not attain conventional levels of significance (effectiveness, $M_{\hat{\rho}} = .232$; efficiency, $M_{\hat{\rho}} = .343$; $Z = 1.386$, $p = .166$).

Team Workflow

The final hypothesis we examined held that cohesion would be beneficial to groups whose patterns of team workflow required greater workflow between members. Following the typology of Tesluk et al. (1997), we compared four progressively increasing patterns of team workflow: pooled, sequential, reciprocal, and intensive. Because the underlying moderator construct is probably best expressed as a continuum ranging from low levels of workflow to high levels of workflow, we placed the four categories on a 4-point continuum and examined how well the continuous moderator predicted the corrected correlations using weighted least squares regression (see Steel & Kammeyer-Mueller, 2002, for more details on the analysis of continuous moderators in meta-analysis).[4] The results of this analysis found that the teamwork moderator accounted for a significant amount of variance in the corrected correlations ($R^2 = .096$), $F(1, 52) = 5.430$, $p < .05$. As predicted, as team workflow increased, the cohesion–performance relation became stronger.

Discussion

Our hypotheses for this meta-analysis mainly were concerned with three issues: the independent contribution of each component of cohesion, an examination of the criterion domain, and the role of team workflow. With respect to the components of cohesion, interpersonal attraction, task commitment, and group pride all displayed independent relations to group performance. As we noted in the Results section, however, our assessment of group pride–performance relations is somewhat limited by the small number of effects available. Moreover, the six effects that measured group pride all came from a rather homogeneous set of studies. All were studies of sports teams that used only effectiveness measures of outcomes (usually a win–loss ratio). Clearly, more research is needed to determine the importance of this component of cohesion.

Our analyses of the criterion domain revealed two noteworthy findings. First, in line with Campbell et al. (1993; Campbell, Gasser, & Oswald, 1996), performance behaviors exhibited stronger relations with cohesion than did performance outcomes. In addition, measures of efficiency reflected the beneficial effects of group cohesion better than measures of effectiveness. Finally, patterns of team workflow moderated the cohesion–performance relation; tasks exhibiting greater amounts of workflow also held stronger cohesion–performance relations.

Relating Components of Cohesion to Various Performance Criteria

We attempted to assess the relations between each component of cohesion and multiple criterion categories. Unfortunately, we could not examine each criterion variable with each component of cohesion because of a lack of studies in one or more of the categories. Interpersonal attraction exhibited correlations that largely were consistent with the overall findings. In contrast, task commitment displayed weaker differences between behavior and outcome, as well as effectiveness and efficiency. One potential explanation for the component differences is that task commitment, in contrast to interpersonal attraction, may not be as likely to reflect the fluidity of group behavior as it is to reflect an overall stronger motivation to perform well (Festinger, Schachter, & Back, 1950). As such, the benefits of task commitment apply despite the particular choices of performance measurement.

It is clear, however, that these propositions are in need of more focused research attention. To our knowledge, no researcher has attempted a direct test of components of cohesion with any of the criteria examined in this meta-analysis. Indeed, this dearth of research contributed to our inability to examine the relations between performance criteria and group pride, as well as the relations between any component and patterns of teamwork.

When Are Cohesive Groups Advantageous?

A consideration of the relations between components of cohesion and various criteria reveals several situations in which cohesive groups are likely to perform better. Certainly, when efficiency is an important goal in the organization (i.e., as opposed to situations in which successful completion of a task is the main requirement, e.g., winning a game, successful surgical performance, obtaining a high grade on a project, etc.), cohesive groups gain an advantage. Of particular interest was the finding that the distinction between efficiency and effectiveness was not as strong when cohesion was conceived of as task commitment. This finding is not to say, however, that task commitment does not bear strong relations to measures of effectiveness or efficiency. In fact, our results suggest that the reverse is true; task commitment held moderate relations across all of the examined criterion domains. Whether this finding holds for patterns of team workflow, however, must await future studies that compare this moderator with the various components of cohesion.

Cohesive groups also achieve performance benefits when group performance is conceptualized as a behavior instead of an outcome. We argued that the stronger relation for performance behavior was due to external impediments inherent in outcome measures; however, we also point out that because performance behaviors are causally antecedent to performance outcomes, they may be more closely tied to cohesion and the group processes that result from cohesion. Indeed, one would have difficulty imagining any part of a cohesion–outcome relationship that is not mediated

[4] Consistent with research examining continuous moderator tests (Steel & Kammeyer-Mueller, 2002), we used inverse sampling error weighting for this analysis. However, we relied on the RBNL estimates of sampling error instead of more simplistic variance formulas that do not take artifact information into account.

by performance behavior. This finding illustrates the necessity of defining performance as behavior in order to appropriately identify those constructs that are predictive. Campbell et al. (1996) lamented that applied psychology would be hindered to the extent that we conceive of performance as an outcome, and this meta-analysis provides justification for their disapprobation.

Perhaps the most interesting point concerning the utility of group cohesion is that every way in which we examined components of cohesion and domains of criteria resulted in a positive mean correlation, and all of these mean correlations were significantly different from zero. Certainly, previous meta-analyses on the overall cohesion–performance effect left little doubt that cohesion benefits performance, but our analysis suggests that this benefit cuts across many different conceptualizations of the cohesion and performance constructs. In addition, although differences in magnitude between the mean corrected correlations were often not large (e.g., effect size difference of .13 between the respective mean corrected correlations for behavior versus outcome criteria), these small differences in correlations at the group level can have substantial practical utility when considered within a decision theoretic utility sense.

As a final note, we acknowledge the potential relation of cohesion to another criterion domain, broadly termed contextual performance. Although there is some debate concerning what actions or behaviors should be considered indicators of contextual performance, as well as the appropriate nomenclature for this construct domain, generally contextual performance includes "behaviors that do not support the technical core itself so much as they support the broader organizational, social, and psychological environment in which the technical core must function" (Borman & Motowidlo, 1993, p. 73). Because contextual performance often includes actions that are helpful to other members of a group, it seems likely that cohesive groups would experience higher levels of contextual performance (LePine, Hanson, Borman, & Motowidlo, 2000).

Indeed, we attempted to assess this relation, and from the few studies that contained appropriate effects, we found a positive correlation between cohesion and contextual performance. Because contextual performance usually occurs at the individual level, however, most of the studies examining this relation did not meet our inclusion criterion of group-level effects. Nevertheless, we feel it worthwhile to mention our cursory examination as an encouragement to other researchers pursuing this topic.

When Are Cohesive Groups Less Advantageous?

Despite the overall rosy picture presented above for the utility of group cohesion, there definitely were circumstances in which cohesive groups provided little help for performance. Obviously, the flip side of the domains mentioned above require acknowledgment: Compared with performance behaviors, performance outcomes do not reflect the advantages of cohesion, and compared with efficiency measures, effectiveness measures ironically are less effective in capturing the benefits of being in a cohesive group. Finally, as our analysis of team workflow demonstrated, groups who engage in fewer exchanges of work (e.g., pooled workflow) do not benefit from cohesion as much as those groups whose workflow is intensive.

As we noted above, the difference between behavior versus outcome and effectiveness versus efficiency held mostly for studies that conceived of cohesion as interpersonal attraction. Although our meta-analysis was unable to identify the reasons why interpersonal attraction held these stronger relations, there are speculative reasons to expect this result. Efficiency often relies on the communication and cooperation of group members. To the extent that interpersonal attraction facilitates these group processes, it will likely lead to more efficient group performance. If the quality of group inputs is not taken into account, as is the case with measures of effectiveness, then interpersonal attraction may be somewhat insensitive as a predictor. Unfortunately, research directly examining these assertions is largely absent in the organizational literature.

The fact that correlations involving pooled team workflow were weak in comparison with more intensive patterns makes intuitive sense when considering the types of tasks involved in these studies. If the members of a team focus on their own individual performances, then many of the benefits of cohesion would have no bearing on the team's performance. Interestingly, it may be the case that task commitment would boost performance in tasks requiring pooled teamwork whereas interpersonal attraction would not. The benefits of task commitment revolve around a shared motivation to do well on a task. Because this motivation does not necessarily require members to work together, tasks requiring pooled teamwork might still benefit from this shared commitment. The same logic applies to group pride: If members of a team are motivated to maintain the esteem of the group, then their motivation may carry them through individual tasks that are then pooled to the team level. Unfortunately, there were not sufficient numbers of studies in our sample to test these hypotheses, but certainly our thoughts merit consideration in future studies of cohesion and performance.

Limitations

There are several possible limitations present in our meta-analysis that warrant discussion. First, there are issues regarding correlations or confounds amongst moderators in other meta-analyses and within our own. A case could be made, for example, that group size is theoretically related to one or more of our included variables (e.g., team size occasionally is considered an input for efficiency measures). If this were the case, it would be unclear which variable was responsible for our observed results.[5] Obviously, we cannot include all previously examined moderators, and to that extent, we must remain cautious with the interpretation of our results. It is possible, for example, that our ideas and results concerning particular patterns of workflow add little beyond Gully et al.'s (1995) notions of task interdependence. We would argue, however, that identifying the manner in which work is exchanged in a group provides a more precise understanding of the actual process involved. Indeed, team workflow describes one of the potential mechanisms that make groups interdependent. At a minimum, conceptualizing workflow as a continuous variable allows for a more precise description of the actual process of teamwork

[5] Because group size was fairly easy to code, we did examine this possibility. We recomputed our average effects for each variable controlling for group size. No meta-analytic effect was significantly different from its original estimate. Furthermore, the deviation of greatest magnitude was rather small (.019).

than do categories of task interdependence, and results of our moderator analyses are consistent with expectations for workflow as a continuous variable.

Aside from overlap with moderators of other meta-analyses, it is possible that variables within our own meta-analyses were confounded. For example, if a significant proportion of the same set of studies included both outcome criteria and measures of effectiveness, then it is difficult to say which variable is responsible for the average effect. Indeed, of the six possible relations between our four variables, which, for the purpose of breaking down distributions of correlations, were treated as moderators (excluding group pride because of the small number of effects), two correlations were significant. Behavior–outcome was related to effectiveness–efficiency ($r = .456$) such that studies with behavioral measures of performance tended also to use efficiency criteria (or, conversely, that studies with measures of outcomes tended also to use effectiveness criteria; see Appendix B, however, for examples of studies that were coded in the four possible combinations of categories). In addition, pattern of team workflow was related to behavior–outcome ($r = .389$) such that studies with more workflow between members tended to use outcome measures (or, conversely, studies with less workflow between members tended to use behavioral measures).

Appropriate evaluation of these relations, however, is a difficult task. Using a weighted least squares regression approach (Steel & Kammeyer-Mueller, 2002), as we did in some other analyses, proved difficult to conduct for reasons of study exclusion. In particular, many studies had to be eliminated because they could not be coded for all relevant variables. Further still, some studies, despite having an effect size estimate for all of the relevant categories, did not have the same effect size estimate. Thus, the remaining set of studies was small in comparison to the total sample of studies, and more importantly, the criteria for elimination was potentially biased. Our solution, therefore, is to encourage future researchers to tackle the question of whether stronger cohesion–performance relations surface as a result of assessing behaviors (as opposed to outcomes) by using measures of efficiency (as opposed to effectiveness) or if these relations benefit from both factors in an additive fashion or reflect an unexamined causal ordering of variables (e.g., see Gonzalez, Burke, Santuzzi, & Bradley, in press; Kirkman & Rosen, 1999).

Conclusion

In summary, we tested several hypotheses in the cohesion–performance literature using meta-analytic techniques. Our results provide compelling evidence for expected differential relations between group cohesion and different types of criteria. Similarly, our sample of studies found that groups who take the most advantage of cohesion typically engage in intensive patterns of workflow. Finally, these results suggest that all three of Festinger's (1950) original components of cohesion—interpersonal attraction, task commitment, and group pride—each bear significant independent relations to performance across many criterion categories. These findings not only enhance our understanding of the construct domains of group cohesion and group performance, but also add to our knowledge of the magnitudes of effects between constructs in the respective domains. A final benefit of our meta-analyses was in identifying important gaps in our understanding of relations between group cohesion and group performance, hopefully providing guidance for future primary studies.

References

References marked with an asterisk indicate studies included in the meta-analyses.

*Arroyo, D. G. (1997). *Cohesion, performance, and satisfaction in the co-acting sport of collegiate wrestling.* Unpublished doctoral dissertation, University of Connecticut, Storrs.

Austin, J. T., & Villanova, P. (1992). The criterion problem: 1917–1992. *Journal of Applied Psychology, 77,* 836–874.

*Bakeman, R., & Helmreich, R. (1975). Cohesiveness and performance: Covariation and causality in an undersea environment. *Journal of Experimental Social Psychology, 11,* 478–489.

*Barrick, M. R., Stewart, G. L., Neubert, M. J., & Mount, M. K. (1998). Relating member ability and personality to work-team processes and team effectiveness. *Journal of Applied Psychology, 83,* 377–391.

*Bhatara, K. (1972). *Effect of cohesiveness on goal setting behavior in the group.* Jabalpur, India: Banarsidas Bhanot.

*Bird, A. M. (1977). Team structure and success as related to cohesiveness and leadership. *Journal of Social Psychology, 103,* 217–223.

*Blades, J. W. (1986). *Rules for leadership: Improving unit performance.* Washington, DC: National Defense University Press.

Borman, W. C., & Motowidlo, S. J. (1993). Expanding the criterion domain to include elements of contextual performance. In N. Schmitt & W. C. Borman (Eds.), *Personnel selection in organizations* (pp. 71–98). San Francisco: Jossey-Bass.

Borucki, C. C., & Burke, M. J. (1999). An examination of service-related antecedents to retail store performance. *Journal of Organizational Behavior, 20,* 943–962.

*Burchfield, M. A. (1997). *Personality composition as it relates to team performance.* Unpublished doctoral dissertation, Stevens Institute of Technology, Hoboken, New Jersey.

Burke, M. J., & Landis, R. S. (2003). Methodological and conceptual challenges in conducting and interpreting meta-analyses. In K. R. Murphy (Ed.), *Validity generalization: A critical review* (pp. 287–309). Mahwah, NJ: Erlbaum.

Campbell, J. P. (1990). Modeling the performance prediction problem in industrial and organizational psychology. In M. D. Dunnette & L. M. Hough (Eds.), *Handbook of industrial and organizational psychology* (2nd ed., Vol. 1, pp. 687–732). Palo Alto, CA: Consulting Psychologists Press.

Campbell, J. P., Gasser, M. B., & Oswald, F. L. (1996). The substantive nature of performance variability. In K. R. Murphy (Ed.), *Individual differences and behavior in organizations* (pp. 258–299). San Francisco: Jossey-Bass.

Campbell, J. P., McCloy, R. A., Oppler, S. H., & Sager, C. E. (1993). A theory of performance. In N. Schmitt & W. C. Borman (Eds.), *Personnel selection in organizations* (pp. 35–70). San Francisco: Jossey-Bass.

Carless, S. A, & De Paola, C. (2000). The measurement of cohesion in work teams. *Small Group Research, 31,* 71–88.

*Carpenter, S., & Radhakrishnan, P. (2002). *The relation between allocentrism and perceptions of in-groups.* Manuscript submitted for publication.

Carron, A. V. (1982). Cohesiveness in sport groups: Interpretations and considerations. *Journal of Sport Psychology, 4,* 123–138.

*Carron, A. V., & Ball, J. R. (1977). An analysis of the cause–effect characteristics of cohesiveness and participation motivation in intercollegiate hockey. *International Review of Sport Sociology, 12,* 49–60.

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., & Brawley, L. (1985). The development of an instrument to assess cohesion in sport teams: The Group Environment Questionnaire. *Journal of Sport Psychology, 7,* 244–266.

Cartwright, D. (1968). The nature of group cohesiveness. In D. Cartwright & A. Zander (Eds.), *Group dynamics: Research and theory* (3rd ed., pp. 91–109). New York: Harper & Row.

*Cohen, D., Whitmyre, J. W., & Funk, W. H. (1960). Effect of group cohesiveness and training upon creative thinking. *Journal of Applied Psychology, 44,* 319–322.

*Colarelli, S. M., & Boos, A. L. (1992). Sociometric and ability-based assignment to work groups: Some implications for personnel selection. *Journal of Organizational Behavior, 13,* 187–196.

Cooper, H. (1998). *Synthesizing research: A guide for literature reviews* (3rd ed.). Thousand Oaks, CA: Sage.

Cornwell, J. M. (1993). Monte Carlo comparisons of three tests for homogeneity of independent correlations. *Educational and Psychological Measurement, 53,* 605–618.

Cota, A. A., Evans, C. R., Dion, K. L., Kilik, L., & Longman, R. S. (1995). The structure of group cohesion. *Personality & Social Psychology Bulletin, 21,* 572–580.

*Cotter, L. L. (1979). *Group cohesiveness and team success among women's intercollegiate basketball teams.* Unpublished doctoral dissertation, Indiana University, Bloomington.

*Craig, T. Y., & Kelly, J. R. (1999). Group cohesiveness and creative performance. *Group Dynamics: Theory, Research, & Practice, 3,* 243–258.

*Darley, J. G., Gross, N., & Martin, W. C. (1952). Studies of group behavior: Factors associated with the productivity of groups. *Journal of Applied Psychology, 36,* 396–403.

Davis, J. (1969). *Group performance.* Reading, MA: Addison-Wesley.

*Deep, S. D., Bass, B. M., & Vaughan, J. A. (1967). Some effects on business gaming of previous quasi-*T* group affiliations. *Journal of Applied Psychology, 51,* 426–431.

*Duffy, M. K., & Shaw, J. D. (2000). The Salieri Syndrome: Consequences of envy in groups. *Small Group Research, 31,* 3–23.

*Eisenberg, J. (2001). *The effects of reward schemes, individualism-collectivism, and intrinsic motivation on teams' creative performance.* Unpublished doctoral dissertation, Colorado State University, Fort Collins.

*Elias, F. G., Johnson, M. E., & Fortman, J. B. (1989). Task-focused self-disclosure: Effects on group cohesiveness, commitment to task, and productivity. *Small Group Behavior, 20,* 87–96.

Evans C. R., & Dion, K. L. (1989). Group cohesion and performance: A meta-analysis. *Small Group Research, 22,* 175–186.

Festinger, L. (1950). Informal social communication. *Psychological Review, 57,* 271–282.

Festinger, L., Schachter, S., & Back, K. (1950). *Social pressures in informal groups: A study of human factors in housing.* New York: Harper & Brothers.

*Fiedler, F. E. (1954). Assumed similarity measures as predictors of team effectiveness. *Journal of Abnormal & Social Psychology, 49,* 381–388.

Finkelstein, L. M., Burke, M. J., & Raju, N. S. (1995). Age discrimination in simulated employment contexts: An integrative analysis. *Journal of Applied Psychology, 80,* 652–663.

Flanagan, J. C. (1956). The evaluation of methods in applied psychology and the problem of criteria. *Occupational Psychology, 30,* 1–9.

*Fodor, E. M., & Smith, T. (1982). The power motive as an influence on group decision making. *Journal of Personality and Social Psychology, 42,* 178–185.

*Fox, E. C. (1986). *Team cohesion, ability, and coaches' leadership effectiveness as predictors of success in women's intercollegiate softball.* Unpublished doctoral dissertation, University of Georgia, Athens.

*Gekoski, N. (1952). Predicting group productivity. *Personnel Psychology, 5,* 281–292.

*George, J. M., & Bettenhausen, K. (1990). Understanding prosocial behavior, sales performance, and turnover. *Journal of Applied Psychology, 75,* 698–709.

Gleser L. J., & Olkin, I. (1994). Stochastically dependent effect sizes. In H. Cooper & L. Hedges (Eds.), *The handbook of research synthesis* (pp. 339–355). New York: Russell Sage Foundation.

*Gonzalez, M. G., Burke, M. J., Santuzzi, A. M., & Bradley, J. (in press). The impact of group process variables on the effectiveness of distance collaboration groups. *Computers in Human Behavior.*

*Goodacre, D. M. (1951). The use of a sociometric test as a predictor of combat unit effectiveness. *Sociometry, 14,* 148–152.

Gosenpud, J. (1989). The prediction of simulation performance as it is affected by time. *Simulation & Games, 20,* 319–350.

Gowda, R. G. (1988). *Influence of individual characteristics and group cohesiveness on programmer productivity.* Unpublished doctoral dissertation, Georgia State University.

*Greene, C. N. (1989). Cohesion and productivity in work groups. *Small Group Behavior, 20,* 70–86.

Gully, S. M., Devine, D. J., & Whitney, D. J. (1995). A meta-analysis of cohesion and performance: Effects of level of analysis and task interdependence. *Small Group Research, 26,* 497–520.

Hackman, J. R. (1968). Effects of task characteristics on group products. *Journal of Experimental Social Psychology, 4,* 162–187.

Hackman, J. R., & Morris, C. G. (1975). Group tasks, group interaction process, and group performance effectiveness: A review and proposed integration. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 8, pp. 45–99). New York: Academic Press.

Harrison, D. A., Price, K. H., & Bell, M. P. (1998). Beyond relational demography: Time and the effects of surface- and deep-level diversity on work group cohesion. *Academy of Management Journal, 41,* 96–107.

*Hemphill, J. K., & Sechrest, L. B. (1952). A comparison of three criteria of aircrew effectiveness in combat over Korea. *Journal of Applied Psychology, 36,* 323–327.

Hollingshead, A. B. (1998). Communication, learning and retrieval in transactive memory systems. *Journal of Experimental Social Psychology, 34,* 423–442.

Hollingshead, A. B. (2000). Perceptions of expertise and transactive memory in work relationships. *Group Processes and Intergroup Relations, 3,* 257–267.

*Hoegl, M., & Gemuenden, H. G. (2001). Teamwork quality and the success of innovative projects: A theoretical concept and empirical evidence. *Organization Science, 12,* 435–449.

Hunter, J. E., & Schmidt, F. L. (1990). *Methods of meta-analysis: Correcting for sources of error and bias in research findings.* Newbury Park, CA: Sage.

*Jaffe, E. D., & Nebenzahl, I. D. (1990). Group interaction and business game performance. *Simulation and Gaming, 21,* 133–146.

*Jehn, K. A. (1994). Enhancing effectiveness: An investigation of advantages and disadvantages of value-based intragroup conflict. *International Journal of Conflict Management, 5,* 223–238.

*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.

Kabanoff, B., & O'Brien, G. E. (1979). The effects of task type and cooperation upon group products and performance. *Organizational Behavior & Human Decision Processes, 23,* 163–181.

*Karau, S. J., & Hart, J. W. (1998). Group cohesiveness and social loafing: Effects of a social interaction manipulation on individual motivation within groups. *Group Dynamics, 2,* 185–191.

*Keller, R. T. (1986). Predictors of the performance of project groups in R & D organizations. *Academy of Management Journal, 29,* 715–726.

*Keyton, J., & Springston, J. (1990). Redefining cohesiveness in groups. *Small Group Research, 21,* 234–254.

Kirchler, E., & Davis, J. H. (1986). The influence of member status differences and task type on group consensus and member position change. *Journal of Personality and Social Psychology, 51,* 83–91.

Kirkman, B. L., & Rosen, B. (1999). Beyond self-management: Antecedents and consequences of team empowerment. *Academy of Management Journal, 42,* 58–74.

*Klein, H. J., & Mulvey, P. W. (1995). Two investigations of the relationships among group goals, goal commitment, cohesion, and performance. *Organizational Behavior and Human Decision Processes, 61,* 44–53.

Klein, K. J., Dansereau, F. & Hall, R. J. (1994). Levels issues in theory development, data collection, and analysis. *Academy of Management Review, 19,* 195–229.

Koslowsky, M., & Sagie, A. (1993). On the efficacy of credibility intervals as indicators of moderator effects in meta-analytic research. *Journal of Organizational Behavior, 14,* 695–699.

*Landers, D. M., Wilkinson, M. O., Hatfield, B. D., & Barber, H. (1982). Causality and the cohesion–performance relationship. *Journal of Sport Psychology, 4,* 170–183.

*Langfred, C. W. (1998). Is group cohesiveness a double-edged sword? An investigation of the effects of cohesiveness on performance. *Small Group Research, 29,* 124–143.

*Langfred, C. W. (2000). The paradox of self-management: Individual and group autonomy in work groups. *Journal of Organizational Behavior, 21,* 563–585.

LePine, J. A., Hanson, M., Borman, W., & Motowidlo, S. J. (2000). Contextual performance and teamwork: Implications for staffing. In G. R. Ferris & K. M. Rowland (Eds.), *Research in personnel and human resources management* (Vol. 19, pp. 53–90). Stamford, CT: JAI Press.

*Lodahl, T. M., & Porter, L. W. (1961). Psychometric score patterns, social characteristics, and productivity of small industrial workgroups. *Journal of Applied Psychology, 45,* 73–79.

*Lorenz, C. E. (1985). *The influence of expectations and cohesiveness on the performance of small groups: A study of quality circles.* Unpublished doctoral dissertation, North Carolina State University, Raleigh.

Lott, A. J., & Lott, B. E. (1965). Group cohesiveness as interpersonal attraction: A review of relationships with antecedent and consequent variables. *Psychological Bulletin, 64,* 259–309.

*Martens, R., & Peterson, J. A. (1971). Group cohesiveness as a determinant of success and member satisfaction in team performance. *International Review of Sport Sociology, 6,* 49–59.

Mathieu, J. E., Heffner, T. S., Goodwin, G. F., Salas, E., & Cannon-Bowers, J. A. (2000). The influence of shared mental models on team process and performance. *Journal of Applied Psychology, 85,* 273–283.

Matheson, H., Mathes, S., & Murray, M. (1996). Group cohesion of female intercollegiate coacting and interacting teams across a competitive season. *International Journal of Sport Psychology, 27,* 37–49.

McGrath, J. E. (1984). *Groups, interaction and performance.* Englewood Cliffs, NJ: Prentice-Hall.

*Melnick, M. J., & Chemers, M. M. (1974). Effects of group social structure on the success of basketball teams. *Research Quarterly, 45,* 1–8.

Mickelson, J. S., & Campbell, J. H. (1975). Information behavior: Groups with varying levels of interpersonal acquaintance. *Organizational Behavior and Human Decision Processes 13,* 193–205.

*Miesing, P., & Preble, J. F. (1985). Group process and performance in a complex business simulation. *Small Group Behavior, 16,* 325–338.

*Mossholder, K. W., & Bedeian, A. G. (1983). Group interaction processes: Individual and group-level effects. *Group & Organization Studies, 8,* 187–202.

Mudrack, P. E. (1989). Group cohesiveness and productivity: A closer look. *Human Relations, 9,* 771–785.

Mullen, B., & Copper, C. (1994). The relation between group cohesiveness and performance: An integration. *Psychological Bulletin, 115,* 210–227.

*Mulvey, P. W., & Klein, H. J. (1998). The impact of perceived loafing and collective efficacy in group goal processes and group performance. *Organizational Behavior and Human Decision Processes, 74,* 62–87.

*Neal, D. J. (1997). Group competitiveness and cohesion in a business simulation. *Simulation & Gaming, 28,* 460–476.

*Neubert, M. J. (1999). Too much of a good thing or the more the merrier? Exploring the dispersion and gender composition of informal leadership in manufacturing teams. *Small Group Research, 30,* 635–646.

*Norris, D. R., & Niebuhr, R. E. (1980). Group variables and gaming success. *Simulation and Games, 11,* 301–312.

Oliver, L. W., Harman, J., Hoover, E., Hayes, S. M., & Pandhi, N. A. (1999). A quantitative integration of the military cohesion literature. *Military Psychology, 11,* 57–83.

Ostroff, C. (1993). Comparing correlations based on individual-level and aggregated data. *Journal of Applied Psychology, 78,* 569–582.

Paese, P. W., & Switzer, F. S., III. (1988). Validity generalization and hypothetical reliability distributions. *Journal of Applied Psychology, 73,* 267–274.

*Podsakoff, P. M., MacKenzie, S. B., & Ahearne, M. (1997). Moderating effects of goal acceptance on the relationship between group cohesiveness and productivity. *Journal of Applied Psychology, 82,* 974–983.

*Porter, T. W., & Lilly, B. S. (1996). The effects of conflict, trust, and task commitment on project team performance. *International Journal of Conflict Management, 7,* 361–376.

Raju, N. S., Burke, M. J., Normand, J., & Langlois, G. M. (1991). A new meta-analytic approach. *Journal of Applied Psychology, 76,* 432–446.

Raju, N. S., & Drasgow, F. (2003). Maximum likelihood estimation in validity generalization. In K. R. Murphy (Ed.), *Validity generalization: A critical review* (pp. 263–286). Hillsdale, NJ: Erlbaum.

Raju, N. S., Pappas, S., & Williams, C. P. (1989). An empirical Monte Carlo test of the accuracy of the correlation, covariance, and regression slope models for assessing validity generalization. *Journal of Applied Psychology, 74,* 901–911.

Sackett, P. R., Harris, M. M., & Orr, J. M. (1986). On seeking moderator variables in the meta-analysis of correlational data: A Monte Carlo investigation of statistical power and resistance to Type-1 error. *Journal of Applied Psychology, 71,* 302–310.

Schmidt, F. L., & Hunter, J. E. (1977). Development of a general solution to the problem of validity generalization. *Journal of Applied Psychology, 62,* 529–540.

Scullen, S. E. (1997). When ratings from one source have been averaged, but ratings from another source have not: Problems and solutions. *Journal of Applied Psychology, 82,* 880–888.

Seashore, S. E. (1954). *Group cohesiveness in the industrial work group.* Ann Arbor: Institute for Social Research, University of Michigan.

*Seers, A., Petty, M. M., & Cashman, J. F. (1995). Team-member exchange under team and traditional management: A naturally occurring quasi-experiment. *Group & Organization Management, 20,* 18–38.

*Smith, K. G., Smith, K. A., Olian, J. D., Sims, H. P., O'Banon, D. P., & Scully, J. A. (1994). Top management team demography and process: The role of social integration and communication. *Administrative Science Quarterly, 39,* 412–438.

Smith, P. C. (1976). Behavior, results, and organizational effectiveness: The problem of criteria. In M. D. Dunnette (Ed.), *Handbook of industrial and organizational psychology* (pp. 745–775). Chicago: Rand McNally.

Spector, P. E., & Levine, E. L. (1987). Meta-analysis for integrating study outcomes: A Monte Carlo study of its susceptibility to Type I and Type II errors. *Journal of Applied Psychology, 72,* 3–9.

Steel, P. D., & Kammeyer-Mueller, J. D. (2002). Comparing meta-analytic moderator estimation techniques under realistic conditions. *Journal of Applied Psychology, 87,* 96–111.

Steiner, I. (1972). *Group processes and productivity.* San Diego, CA: Academic Press.

Stewart, G. L., & Barrick, M. R. (2000). Team structure and performance:

Assessing the mediating role of intrateam process and the moderating role of task type. *Academy of Management Journal, 43,* 135–148.
*Stinson, J. E., & Hellebrandt, E. T. (1972). Group cohesiveness, productivity, and strength of formal leadership. *Journal of Social Psychology, 87,* 99–105.
Stogdill, R. M. (1972). Group productivity, drive, and cohesiveness. *Organizational Behavior and Human Performance, 8,* 26–43.
*Tehan, T. A. (1983). *The development and evaluation of effectiveness of cohesiveness training with high school basketball teams.* Unpublished doctoral dissertation, Virginia Commonwealth University, Richmond.
*Terborg, J. R., Castore, C., & DeNinno, J. A. (1976). A longitudinal field investigation of the impact of group composition on group performance and cohesion. *Journal of Personality and Social Psychology, 34,* 782–790.
*Tesluk, P. E., & Mathieu, J. E. (1999). Overcoming roadblocks to effectiveness: Incorporating management of performance barriers into models of work group effectiveness. *Journal of Applied Psychology, 84,* 200–217.
Tesluk, P. E., Mathieu, J. E., Zaccaro, S. J., & Marks, M. A. (1997). Task and aggregation issues in the analysis and assessment of team performance. In M. T. Brannick, E. Salas, & C. Prince (Eds.), *Team performance and measurement: Theory, methods, and applications* (pp. 197–224). Mahwah, NJ: Erlbaum.
Thompson, J. D. (1967) *Organizations in action.* New York: McGraw-Hill.
Tziner, A. (1982). Differential effects of group cohesiveness types: A clarifying overview. *Social Behavior and Personality, 10,* 227–239.
*Tziner, A., & Vardi, Y. (1983). Ability as a moderator between cohesiveness and tank crews performance. *Journal of Occupational Behavior, 4,* 137–143.
Wanous, J. P., Sullivan, S. E., & Malinak, J. (1989). The role of judgment calls in meta-analysis. *Journal of Applied Psychology, 74,* 259–264.
*Wech, B. A., Mossholder, K. W., Steel, R. P., & Bennett, N. (1998). Does work group cohesiveness affect individuals' performance and organizational commitment? A cross-level examination. *Small Group Research, 29,* 472–494.
Weed, S. E., Mitchell, T. R., & Moffitt, W. (1976). Leadership style, subordinate personality, and task type as predictors of performance and satisfaction with supervision. *Journal of Applied Psychology, 61,* 58–66.
Wegner, D. M., Erber, R., & Raymond, P. (1991). Transactive memory in close relationships. *Journal of Personality and Social Psychology, 61,* 923–929.
*Wekselberg, V., Goggin, W. C., & Collings, T. J. (1997). A multifaceted concept of group maturity and its measurement and relationship to group performance. *Small Group Research, 28,* 3–28.
Whitener, E. M. (1990). Confusion of confidence intervals and credibility intervals in meta-analysis. *Journal of Applied Psychology, 75,* 315–321.
Wilderom, P. M., Glunk, U., & Maslowski, R. (2000). Organizational culture as a predictor of organizational performance. In N. M. Ashkanasy, C. P. M. Wilderom, & M. F. Peterson (Eds.), *Handbook of organizational culture and climate* (pp. 193–209). Thousand Oaks, CA: Sage.
*Williams, J. M., & Hacker, C. M. (1982). Causal relationships among cohesion, satisfaction, and performance in women's intercollegiate field hockey teams. *Journal of Sport Psychology, 4,* 324–337.
*Wolfe, J., & Box, T. M. (1988). Team cohesion effects on business game performance. *Simulation and Games, 19,* 82–98.
*Wong, L. (1992). *The effects of cohesion on organizational performance: A test of two models.* Unpublished doctoral dissertation, Texas Technological University, Lubbock.
*Zaccaro, S. J., Gualtieri, J., & Minionis, D. (1995). Task cohesion as a facilitator of team decision making under temporal urgency. *Military Psychology, 7,* 77–93.
*Zaccaro, S. J., & Lowe, C. A. (1986). Cohesiveness and performance on an additive task: Evidence for multidimensionality. *Journal of Social Psychology, 128,* 547–558.
*Zaccaro, S. J., & McCoy, C. (1988). The effects of task and interpersonal cohesiveness on performance of a disjunctive group task. *Journal of Applied Social Psychology, 18,* 837–851.

(*Appendixes follow*)

Appendix A

Examples and Coding Agreement for Each Moderator Variable Included in the Study

Characteristic	Examples	Agreement (κ)[a]
Type of performance		.84
Behavior	Ratings of specific combat behaviors during tactical field problems, supervisory ratings of specific work behaviors (e.g., knowledge of tasks, planning, and so forth), cognitive performance on puzzle tasks	
Outcome	Final grades on class projects, supervisor ratings of work group outcomes, and win–loss ratio in sports teams	
Type of performance measure		.89
Effectiveness	Ratings or counts of total group output, team ranking in a business simulation, and quality of group decisions	
Efficiency	Group output over a specific period of time, return on investment in simulation games and actual organizations, brainstorming results over a specific period of time, and supervisor ratings of efficiency	
Pattern of team workflow		.96
Pooled	Individual sales performance (that was then aggregated to the group level), solving of individual puzzles, and collegiate wrestling competitions	
Sequential	Clerical work that proceeded in stages and group card-sorting tasks (in which one member follows the same person each time)	
Reciprocal	Surveying teams, group puzzle tasks (in which members interact with other group members one at a time to complete the task), and class projects conducted over e-mail (i.e., members interact with other members, but there is no simultaneous interaction)	
Intensive	Decision-making tasks, group puzzle tasks (in which members must interact with the other team members simultaneously to complete the puzzle), business simulations, and class group projects.	
Component of cohesion		
Interpersonal attraction	Manipulations of attitude similarity, sociometric nominations, measures of attitude similarity, and any items indicating preference for the members of the group	
Task commitment	Measures and manipulations of task enjoyment, importance, or attraction (personal or group)	
Group pride	Items assessing the value of group membership, importance of belonging to the group, and measures of attraction to the group itself (i.e., apart from its members)	

[a] Two coders rated each category of moderator for all studies. For example, each correlation was examined for whether performance was a behavior, an outcome, or ambiguous with respect to this category.

Appendix B

Examples of Studies That Were Multiply Coded Under the Behavior–Outcome and Effectiveness–Efficiency Categories

	Effectiveness	Efficiency
Behavior	Total time spent engaging in work-related activities (Bakeman & Helmreich, 1975); the number of action proposals in a group decision-making task (Fodor & Smith, 1982); overall evaluation of behaviors exhibited in an oral presentation (Keyton & Springston, 1990)	Number of ideas generated in twelve minutes for brainstorming-task (Cohen et al., 1960); summing and extending invoices in a seven minute period (Wong, 1992); number of product solution ideas, weighted by ratings of creativity (Eisenberg, 2001)
Outcome	Win–loss ratio for athletic teams (Bird, 1977; Carron & Ball, 1977); overall score on Moon Survival Task (Carpenter & Radhakrishnan, 2002); rank or overall grade on a class project or business simulation (Colarelli & Boos, 1992; Neal, 1997)	Points gained in a fifteen minute period on a cargo-loading simulation (Burchfield, 1997); supervisory rating of units produced and errors made (Gekoski, 1952); total monthly sales, divided by number of employees (George & Bettenhausen, 1990)

Received January 7, 2002
Revision received January 28, 2003
Accepted March 24, 2003 ■