



A Quantitative Integration of the Military Cohesion Literature

Laurel W. Oliver, Joan Harman, Elizabeth Hoover, Stephanie M. Hayes & Nancy A. Pandhi

To cite this article: Laurel W. Oliver, Joan Harman, Elizabeth Hoover, Stephanie M. Hayes & Nancy A. Pandhi (1999) A Quantitative Integration of the Military Cohesion Literature, *Military Psychology*, 11:1, 57-83, DOI: [10.1207/s15327876mp1101_4](https://doi.org/10.1207/s15327876mp1101_4)

To link to this article: https://doi.org/10.1207/s15327876mp1101_4



Published online: 17 Nov 2009.



Submit your article to this journal [↗](#)



Article views: 83



View related articles [↗](#)



Citing articles: 30 View citing articles [↗](#)

A Quantitative Integration of the Military Cohesion Literature

Laurel W. Oliver, Joan Harman, Elizabeth Hoover,
Stephanie M. Hayes, and Nancy A. Pandhi

*U.S. Army Research Institute
Alexandria, Virginia*

In this article, we review research on group cohesion in military units, using meta-analytic techniques that convert research findings to a common metric. After combining some of the 39 samples to eliminate possible overlap in participants, we conducted separate analyses for 7 different outcomes. Effect sizes (correlation coefficients), weighted by the number of participants, were .40 for cohesion and group performance, .20 for cohesion and individual performance, and .47 for cohesion and job/military satisfaction. Cohesion was also positively related to measures of retention (.22), well-being (.24), and readiness (.30) and was inversely related to rates of indiscipline (-.14), although these latter results are not definitive because of the smaller number of samples involved. Effect sizes weighted by the number of groups generally tended to decrease the number of cases, increase the effect sizes, and broaden the confidence intervals. We conclude that group cohesion results in desirable outcomes for the military and that the research has implications for policy and training issues. We recommend that future cohesion researchers include all basic data in their research reports, delineate their concepts explicitly, use valid and reliable measures, and give careful consideration to the level of analysis and other moderator variables.

Group cohesion has been the subject of considerable research over the past several decades. Although cohesion is generally viewed as a desirable characteristic of groups and one that is positively related to productivity and performance, the extensive research on the topic has heretofore resulted—as is often the case—in conflicting findings. In recent years, meta-analytic strategies (which enable researchers to combine research results quantitatively) have been employed in the integration of the cohesion-performance literature.

Oliver's (1988a) meta-analysis, integrating the results of 14 real-world work groups (of which 9 were military), found a positive relation ($r = .32$) between cohesion and performance. The Evans and Dion (1991) meta-analysis, which included 16 published studies (5 dealing with military groups) and was based on group-level analyses, also found a positive relation between cohesion and performance ($r = .36$ when uncorrected for error of measurement, with $r = .42$ when corrected). Another meta-analysis, conducted by Mullen and Copper (1994), included 49 studies and reported a correlation of .23 for their subset of 10 "military" studies. The latest meta-analysis of cohesion and performance, conducted by Gully, Whitney, and Devine (1995), used 46 studies (5 of which were military) and obtained an uncorrected (for error of measurement) effect size (r) of .166, with a corrected value of .199. Removing one methodologically questionable military study¹ and six effect sizes that used self-report measures of productivity resulted in effect sizes of .221 (uncorrected) and .265 (corrected).

The military believes that cohesive groups enhance the combat effectiveness (e.g., Oliver, 1990). Because there are rarely, if ever, measures of combat effectiveness available, researchers generally employ measures of combat readiness, such as Army General Testing and Evaluation Program scores or ratings by commanders. Although the meta-analyses summarized earlier included military groups, none of them exhaustively sampled that population of studies. Many military studies are not in the refereed literature because they are institutional reports or papers presented at conferences. Hence, a thorough meta-analytic examination of group cohesion in military units has not yet been accomplished, and meta-analyses dealing with group cohesion have not investigated outcomes other than performance.

The purpose of this research was to conduct a meta-analysis of the cohesion research involving military groups and to focus on outcomes valued by the military. The research questions to be explored were as follows:

1. What is the overall relation between group cohesion in military units and various outcomes of interest to the military?
2. Do these cohesion–outcome relations hold across various types of military units? That is, do military units from different branches, different countries, and so on, vary in the degree to which cohesion is related to various outcomes?
3. To what extent are different measures of group cohesion comparable across studies involving military units? Are different cohesion measures conceptually similar? Do different measures of military group cohesion yield essentially the same results, or does operationalizing the construct differently—in terms of formats, number of items, or sources of ratings—lead to different findings?

¹We did not include this study (Hoiberg & Pugh, 1978) in our meta-analysis because the cohesion measure was the respondent's expectations for cohesion at the respondent's next duty station. We felt that expectations did not meet our definition of cohesion, as we have interpreted the construct.

4. How do the results of this research compare to the results of previous meta-analyses of group cohesion?

METHOD

Identification of Pertinent Studies

The original nine military cohesion studies in Oliver's (1988a) meta-analysis were augmented by research that had been reported since that meta-analysis was conducted and by studies that were missed in the original search. We included both published and unpublished research. To identify pertinent studies, we used computer searches, previous reviews, reference lists of studies already identified, a notice in the American Psychological Association Division of Military Psychology newsletter, and personal contacts with scholars known to have conducted research on military cohesion.

Definitions. Many researchers investigating cohesion do not explain what they mean by the constructs they are investigating. What follows are clarifications of the terms *cohesion*, *group*, *outcome*, and *effect size*.

In discussing conceptual issues involved in measuring *cohesion*, Oliver (1990) noted that cohesion is usually considered a "group" variable and conceptualized in some way as the "stick-togetherness" of a group. Although there are many definitions of the cohesiveness construct in the literature, most authorities have agreed that cohesion is a multidimensional construct that may include an interpersonal aspect and a task orientation aspect.

Carron (1982) defined group cohesiveness as "a dynamic process that is reflected in the tendency for a group to stick together and remain united in the pursuit of its goals and objectives" (p. 124). We adopted this statement as our working definition for cohesion.

Another term that needed to be clarified was *group*. For this construct, we accepted a definition proposed by Hogg (1992): "A group of two or more individuals in face-to-face interactions, each aware of the others who belong to the group, and each aware of their positive interdependence as they strive to achieve mutual goals" (p. 4).

An *outcome* is generally regarded as a consequence. Here we define outcome as a consequence or result (associated with military units) that is of interest to the military—for example, combat readiness. Chronologically, cohesion usually precedes outcome, but we do not hold that cohesion causes the outcome, merely that the two may be associated.

Our *effect size* for this meta-analysis was the correlation coefficient (Pearson's r or ρ), a measure of association. This statistic was usually found in our studies.

In a few instances, we had to convert data or another statistic, such as chi-square, to r . Another frequently used effect size, Cohen's d , is often employed in meta-analyses involving experimental manipulations. One can convert either r or d to the other statistic. We found no military cohesion studies with experimental and control groups.

Criteria. We used the following criteria for selecting studies to include in the meta-analysis:

1. The sample consisted of military personnel only. The participants could be students, but they had to be in training for a military career. Groups including both military and civilian personnel did not qualify unless data were reported for each group.
2. The groups were work groups with a mission or task to perform, as contrasted with therapy groups or social groups comprising military personnel.
3. The groups were real groups already in existence, not groups formed for the purpose of the research.
4. The study employed a measure of group cohesion that, in our judgment, conformed reasonably well to our definition of cohesion.
5. The study reported a quantitative measure of the relation between group cohesion and an outcome, such as retention, recognized by the military.

Number of studies located. As a result of our efforts to locate research on military cohesion, we identified 122 documents that we believed concerned cohesion in military units. We found that 40 of these documents met our inclusion criteria and that 82 did not. The 40 we accepted for our meta-analysis reported on 36 different research efforts involving 39 diverse samples. The 82 documents we rejected were not accepted for various reasons: No data or insufficient data were reported in the study, the research lacked a cohesion measure meeting our definition, or the sample included nonmilitary group members.

Coding of Variables

Code sheet. We developed a code sheet incorporating the following variables: the reference (date of the document, source, relation to other studies, etc.), the sample (total number of participants, number of groups, type of sample, etc.), the cohesion measure, the outcome measure(s), other variables investigated in the study, and effect size data for each cohesion–outcome relation.

Coding manual. The purpose of the coding manual was to provide a general orientation to the meta-analysis and to explain details of the coding procedures to

be followed. The guidance to the coders on how to code the variables was intended to be as specific as possible to reduce coding uncertainties.

Coder training and coding procedure. Each study was coded for pertinent variables by two coders: (a) a research psychologist, who had conducted cohesion research and had experience in quantitative research integration; and (b) an advanced graduate student, who had participated in the analysis of military cohesion data.

Coders coded one study together, discussing various aspects of the procedure as they went through the coding process. They then coded three studies independently and met again to discuss their results. Coders coded the remainder of the studies independently, although they occasionally conferred about ambiguous situations. These decisions usually involved which variable to code as cohesion (which sometimes was a subset of another variable, such as morale), which variables to use as outcomes, and when to combine variables. After coding was completed, we reviewed the database to ensure that the data had been entered accurately. All discrepancies were resolved by discussions between the two coders.

Ratings of study quality. We felt that it was important to ascertain the quality of each study in an objective manner. Accordingly, we constructed the Quality of Study Scale, which was based on six coded variables: publication form, reporting of the reliability of the cohesion measure, standardization of the cohesion measure, percentage of outcomes for which reliability was reported, percentage of outcomes that were standardized measures, and percentage of outcomes from the same source as the cohesion measure (reverse coded). Each of the six ratings was on a 3-point scale (coded 1, 2, and 3, from the lowest to highest degree of rigor or quality), with a possible summed range of 6.0 to 18.0. The Quality of Study Scale ratings ranged from 7.0 to 16.00, with a mean across the 39 samples of 11.02 and a standard deviation of 2.64. In addition to this objective measure, the coders made a subjective overall rating of the quality of each research effort on a 5-point scale ranging from 1 (*very poor*) to 5 (*very good*). The coders' ratings were averaged for each study or sample. The range of these averaged subjective ratings was 2.0 to 4.5, with a mean of 3.13 and a standard deviation of 0.65. The subjective overall ratings by the coders correlated .43 with the more objective Quality of Study Scale. Coders rated their confidence in their subjective ratings using a 3-point scale developed by Orwin and Cordray (1985), in which 3 was *certain or almost certain*, 2 was *more likely than not*, and 1 was *guess*. The mean confidence rating for the two coders on this variable was 2.0.

Analyses

As noted by many authorities, the lack of statistical independence among effect sizes may pose a problem for meta-analytic research integration. Lack of independ-

ence may result from using the same sample to calculate more than one effect size from different measures, different samples in the same study may be used to calculate separate effect sizes for each sample, and the same researchers may conduct several different studies (Matt & Cook, 1994). In our analyses, we followed two procedures to ensure greater independence of our data: (a) We combined samples for which independence seemed a problem, and (b) we conducted separate analyses for each class of outcome measure.

Samples and composites. Effect size analyses were based on 33 samples or composite samples. We combined the results from the two studies by Goodacre (1951, 1953), and we combined the studies conducted by Siebold and his colleagues (Julien & Siebold, 1989; Siebold, 1990; Siebold & Kelly, 1987, 1988a, 1988b). The Gal and Manning (1984, 1987) research included military units from two countries (United States and Israel) that were stationed in three different countries, and we treated these three groups as independent samples.

Combining outcomes. We clustered the 21 types of outcome measures enumerated on the code sheet into seven classes of outcomes. Within each study, we averaged effect sizes for outcomes falling into each of the seven different categories. The outcome categories were individual performance (individual scores on a test or an activity such as rifle practice), group performance (group as a whole rated on some measure of performance), well-being (physical or psychological), job/military satisfaction, retention, readiness (operational or psychological), and indiscipline (e.g., AWOL rates). Then, we conducted a separate analysis for each of the seven outcome classes.

Weighting of effect sizes. Typically, effect sizes are weighted by the number of participants involved.² Because cohesion is a group-level construct, however, one can argue that the proper sample is the number of groups. We report the results of both weighting procedures.

Corrections. The *r*-to-*z* transformation is controversial, and we did not use it. Some meta-analysts recommend its use before statistical manipulations involving

²This procedure is in contrast to using the reciprocal of the sampling variance of *r* as recommended by Shadish and Haddock (1994) in their chapter on combining estimates of effect size. Some authorities believe using sample size is preferable to using the variance of *r*, although the former procedure is technically more approximate. B. J. Becker (personal communication, April 3, 1996) has found that, in practice, sample size works better. We found that using sample size resulted in more conservative (i.e., broader) estimates of confidence intervals.

correlation coefficients (e.g., Rosenthal, 1994). Other authorities (e.g., Hunter & Schmidt, 1990) argue against using the transformation because its bias is greater than the bias of the uncorrected r . We also did not correct for error of measurement, as some meta-analysts do (e.g., Gully et al., 1995). Because we found that the reliabilities of instruments were reported poorly in the studies whose results we were integrating, we felt that correcting for measurement error was problematic. By not applying the correction, our mean effect sizes were lower and, thus, more conservative estimates of the true effect sizes.

RESULTS

Description of the Total Sample

Table 1 summarizes the studies we included in our meta-analysis, including descriptions of the 39 samples. Table 2 summarizes the characteristics of these samples.

Date and form of publication. The 39 individual samples spanned a period of 40 years (1951–1991), although almost two thirds (25) were from the 1980s. Some 17 studies (44%) were from published sources (journal articles, books, or a test manual), and another 18 (46%) were conference papers, institutional reports, or dissertations. The 4 remaining studies were miscellaneous unpublished papers.

Number of participants. The number of participants can be reported in different ways—the number of participants originally present in the military units involved, the response rate to a survey, or the number of participants left after those with missing data are deleted. We tried to use samples associated with the analyses, but it was sometimes necessary to estimate or average to obtain a total sample.

In this meta-analysis, the mean number of participants per study or sample was about 955. However, this average is misleading because the range was from 24 to 6,724, with almost two thirds (25) of the samples under 1,000. The total number of participants across all 39 samples was 37,226.

Number and level of groups. We sometimes found it very difficult to determine the number of groups in the study. In 30 of the 39 cases, the authors reported the number of groups, or the coders estimated the number. In the other 9 cases, there was not enough information to make estimates. The number of groups ranged from 4 to 115, with a mean of 34.6 and a median of 24.5. The most frequently reported group level was squad, section, or platoon, which accounted for 12

TABLE 1
Summary Table of Individual Studies and Samples

Study	Sample	Measures	Results
Allen (1986) Allen & Potter (n.d.)	Coast Guard personnel on 12 ships, $N = 615$ (avg)	Cohesion: 3 factors from Cutter Field Survey: Team (12 items), Petty Officers (3 items), Mates (3 items) Outcome 1: Officer judges rated ships on "field unit performance" (7-point scale in 11 areas)	(1) $r = .37$
Blades (1976) Blades (1986)	Army cooks and mess stewards, 45 groups, $N = 225$ (est)	Outcome 2: 5-item work satisfaction scale (work factor) Cohesion: Group Atmospheric score (Fiedler, 1967; 10-item semantic differential)	(2) $r = .46$
Cohen (1957)	Army enlisted personnel in 112 squads, $N = 1,008$	Outcome 1: Mess hall performance rated by Brigade Food Advisors and company commanders on 12 7-point scales Cohesion: Reference group measure (8 items tapping attitudes toward other squad members, importance of squad, and role as squad member)	(1) $r = .11$
Dickenson (1991)	Royal Military College of Canada recruits, 8 squads, $N = 130$	Outcome 1: Squad performance on 3 firing field tests rated by pairs of trained umpires Cohesion: Measure based on Platoon Cohesion Index, adapted officer cadets at College Militaire Royale (English and French)	(1) $r = .16$
Dobbins & Zaccaro (1986)	Members of university Cadets Corps, 10 units (9 squadrons plus battalion staff), $N = 203$ respondents	Outcome 1: Obstacle course performance Outcome 2: Road race performance Outcome 3: Cross-country race performance Outcome 4: Physical fitness test Cohesion: 8-item measure from various sources Outcome 1: Work Satisfaction (5 items from Taylor & Bowers, 1972)	(1) $p = .48$ (2) $p = .05$ (3) $p = .05$ (4) $p = .02$ (1) $r = .41$

TABLE 1 (Continued)

<i>Study</i>	<i>Sample</i>	<i>Measures</i>	<i>Results</i>
Griffith (1986) Griffith (1988)	Matched unit replacement and individual replacement units, 93 companies, $N = 5,366$	Cohesion: Concerned Leadership and Unit Social Climate Survey Scales Outcome 1: General Well-Being Scale (Dupuy, 1978) and Life Satisfaction Scale Outcome 2: Number of doctor visits during previous years Outcome 3: Army satisfaction scale Outcome 4: Reenlistment intent Outcome 5: AWOL and nonjudicial Cohesion: Same scale as for Griffith (1986) Outcome 1: Percentage of soldiers in company attaining "Expert" level of marksmanship	(1) $r = .34$ (2) $r = .07$ (3) $r = .59$ (4) $r = .28$ (5) $r = .06$ (1) $r = .30$
Griffith (1987)	Subsample of Griffith (1986) sample, 43 companies, $N = 4,101$		
Hemphill & Sechrest (1952)	Combat aircrews, 90 groups, $N = 810$ (est)	Cohesion: Sociometric choices for crew positions Outcome 1: Combat performance rated by superiors on 11 scales	(1) $r = .17$
Hoiberg & Berry (1978)	Navy enlisted men, $N = 6,748$ (avg)	Cohesion: 3 scales from Relationships dimension of Military Company Environment Inventory Outcome 1: Navy school success Outcome 2: 2-year active duty survival	(1) $r = .07$ (2) $r = .18$
Hom & Hulin (1981)	Army National Guard personnel from 29 units, $N = 230$	Cohesion: 3 scales from Relationships dimension of Military Company Environment Inventory Outcome 1: Reenlistment	(1) $r = .25$ (see their Table 5)
Julien & Siebold (1989)	Personnel from Army battalions participating in Joint Readiness Training Center exercise, 15 companies, $N = 58$	Cohesion: Platoon Cohesion Index items adapted for Joint Readiness Training Center participation Outcome 1: Platoon-level leaders assessment of typical performance of their squads or platoons	(1) $r = .40$

Manning & Fullerton (1988)	Special Forces support troops (2 battalions), 1 airborne infantry battalion, and randomly selected members of mechanized divisions, $N = 946$	Cohesion: Perceived Social Support (similar to Procidano & Heller, 1983) and author-developed Command Climate Survey Outcome 1: Dupuy's General Well-Being Schedule (see Neff & Husani, 1980) and a marital satisfaction index based on a scale by Roach, Frazier, & Bowden (1981) Outcome 2: Health Perceptions Inventory (Ware, Johnson, Davies-Avery, & Brook, 1979) Outcome 3: Army Satisfaction Inventory (Datel, 1978) Cohesion: Answer to Inspector General questions authors judged measures of cohesion Outcome 1: Army General Testing and Evaluation Program and Operational Readiness Test performance scores Outcome 2: Skill Qualification Test and Physical Training performance scores Outcome 3: Number of administrative discharges, nonjudicial/court-imposed punishments, apprehensions for crimes Outcome 4: Reenlistment Cohesion: Peer cohesion scale of Military Environment Inventory (Moos, 1986) Outcome 1: Companies rank order on performance measures (firing, physical training, graded test scores) Cohesion: Teamwork scale from platoon morale behavioral ratings Outcome 1: Reenlistment Outcome 2: Dissent, drug abuse, and destruction/sabotage Outcome 3: Unit effectiveness	(1) $r = .26$ (2) $r = .25$ (3) $r = .48$ (1) $p = .40$ (2) $p = .52$ (3) $p = .40$ (4) $p = .32$ (1) $p = .58$ (1) $r = .18$ (2) $r = .51$ (3) $r = .79$
Manning & Ingraham (1987)	Stratified sample of 37 persons in each of 20 Army battalions visited by the Inspector General during 9-month period, $N = 740$		
Moos (1986)	Army basic training recruits, 8 companies, $N = 1,432$		
Motowido & Borman (1977)	Personnel from 16 U.S. Army companies, 47 "platoon-sized" units, $N = 753$ (avg)		

(continued)

TABLE 1 (Continued)

<i>Study</i>	<i>Sample</i>	<i>Measures</i>	<i>Results</i>
Nelson & Berry (1968) ^b	Random sample from 191 Marine Corps recruit training platoons, 24 platoons, $N = 1,536$ (avg)	Cohesion: Cohesion index derived from "peer nomination of friendship" Outcome 1: Physical proficiency test and rifle qualifications exercise	(1) $r = .00$
Olson & Borman (1987)	Army project: Personnel selected to represent certain Military Occupational Specialties, $N = 664$ (avg)	Cohesion: Cohesion-cooperation scale of Army Work Environment Questionnaire Outcome 1: Army-wide task ratings, Military Occupational Specialty task rating, soldier effectiveness, hands-on test, and job-knowledge test	(1) $p = .09$
Pleban, Valentine, & Thompson (1987) ^c	Enrollees in Army Advanced Land Reconnaissance Course 46-man teams, $N = 24$	Cohesion: Inclusion scale of Fundamental Interpersonal Relations Orientation-Behavior (Shutz, 1958) Outcome 1: Effectiveness item on Subject Debriefing Questionnaire	(1) $p = 1.0$
Royle (1985)	Marine Corps enlisted women, $N = 574$	Cohesion: Peer Cohesion Scale of Military Company Environment Inventory (Moos, 1973) Outcome 1: Attrition Outcome 2: Stress (perceptions of symptoms) Cohesion: Self-ratings of social support obtained in interviews	(1) $p = .05$ (2) $p = .30$
Shiron (1976)	Israeli Defence Force enlisted personnel in "several" infantry rifle platoons in regiment stationed in Golan Heights, $N = 101$	Outcome 1: Combat performance assessed by peer ratings	(1) $r = .29$
Siebold (1990)	Platoon members in 5 light infantry battalions, 60 platoons, $N = 1,080$	Outcome 2: Soldiers' "feelings" about unit combat preparedness and combat preparedness Cohesion: Platoon Cohesion Index (Siebold & Kelly, 1988a) Outcome 1: Ratings by observer/controllers of platoon performance at Joint Readiness Training Center Outcome 2: Job satisfaction	(2) $r = -.03$ (1) $r = .31$ (2) $r = .70$

Siebold & Kelly (1987)	<p><i>Sample 1:</i> Company members in 2 light infantry battalions, 5 companies, $N = 450$</p> <p><i>Sample 2:</i> Platoon members in 7 light infantry companies, 28 platoons, $N = 504$</p>	<p>Cohesion: Platoon Cohesion Index</p> <p>Outcome 1: Company performance (3 scales) on Road March exercise by "impartial officers from the brigade"</p> <p>Cohesion: Platoon Cohesion Index</p> <p>Outcome 1: Platoon ability to withstand stress rated by company-level leaders</p> <p>Outcome 2: Platoon performance in the field rated by company-level leaders</p>	<p>(1) $r = .84$</p> <p>(1) $r = .60$</p> <p>(2) $r = .55$</p>
Siebold & Kelly (1988a)	Platoon members in 11 light infantry, companies, 44 platoons, $N = 752$	<p>Cohesion: Platoon Cohesion Index</p> <p>Outcome 1: Ratings of platoon performance (garrison/field) by company-level leaders</p> <p>Outcome 2: Ratings by platoon ability to withstand stress by company-level leaders</p> <p>Outcome 3: Ratings of platoon readiness for combat by company-level leaders</p> <p>Outcome 4: Ratings of platoon performance at Joint Readiness Training Center by observer/controllers</p> <p>Cohesion: Platoon Cohesion Index</p> <p>Outcome 1: Rating of platoon performance at Joint Readiness Training Center by observer/controllers and company commanders (avg)</p>	<p>(1) $r = .49$</p> <p>(2) $r = .47$</p> <p>(3) $r = .32$</p> <p>(4) $r = .56$</p> <p>(1) $r = .42$</p>
Siebold & Kelly (1988b)	Platoon members in 3 light infantry companies, 9 platoons, $N = 188$		
Solomon, Mikulincer, & Hobfoll (1986)	Israeli combat soldiers who received treatment at mental health facilities during Israeli-Lebanon War and matched group with no symptoms, $N = 716$	<p>Cohesion: Perception of social support (adapted from Military Company Environment Inventory)</p> <p>Outcome 1: Combat stress reaction (received mental health treatment)</p>	(1) $r = .11$
Sterling & Williams (1982)	U.S. Army personnel (97% E4 or under) stationed at 4 locations in Germany, $N = 399$	<p>Cohesion: Survey items on 3 dimensions (hierarchical, peer cohesion, organizational commitment)</p> <p>Outcome 1: Career intentions</p> <p>Outcome 2: Army satisfaction</p> <p>Outcome 3: Sick call</p>	<p>(1) $r = .23$</p> <p>(2) $r = .29$</p> <p>(3) $r = .07$</p>

(continued)

TABLE 1 (Continued)

Study	Sample	Measures	Results
Turney & Cohen (1976)	Army 16-man communications processing teams (1 NCO, 1 analyst, 14 operators), 9 groups, $N = 117$	Cohesion: Work Environment Questionnaire item rating how well group works together Outcome 1: Work Environment Questionnaire items regarding extent to which fellow operators encourage or emphasize superior performance	(1) $r = .42$
Tziner & Vardi (1983)	Israeli Defence Force tank unit personnel, 115 3-man tank teams, $N = 345$	Cohesion: Sociometric procedure in which "intensity of reciprocity in selection" was measure of cohesiveness Outcome 1: Performance effectiveness	(1) $r = .32$
Vickers, Wallick, & Hervig (1982)	U.S. Marine Corps recruits, 413 randomly selected from 32 basic training battalions and 425 volunteers from 39 basic training battalions, $N = 838$	Cohesion: Questionnaire items for Group Teamwork and Group Support Outcome 1: Satisfaction with Marine Corps	(1) $r = .23$
Yagil (1995)	Israeli Defence Force soldiers and commanders from 18 infantry platoons and 7 armor companies, $N = 547$	Cohesion: Hebrew translation of Platoon Cohesion Index Outcome 1: Unit effectiveness (5 self-report items) Outcome 2: Stress (5 self-report items)	(1) $r = .77$ (2) $r = .63$
Yoest & Tremble (1984)	U.S. Army first-term combat arms soldier, 39 companies, $N = 2,274$	Cohesion: Questionnaire items tapping resistance to leaving unit, quality of work relationships, and interpersonal closeness of soldiers within units Outcome 1: Unit effectiveness Outcome 2: Satisfaction with supervisor	(1) $r = .40$ (2) $r = .24$
Zaccaro (1991)	Members of cadet corps at U.S. university, 9 squadrons, $N = 339$ (avg)	Cohesion: Questionnaire items tapping task and interpersonal cohesion Outcome 1: Demerits for poor performance Outcome 2: Group performance using 7-item index from Survey of Organizations (Taylor & Bowers, 1972) Outcome 3: Absenteeism	(1) $r = .07$ (2) $r = .44$ (3) $r = .11$

Note. avg = average; est = estimate. Negative outcomes reverse coded.

*Gal & Manning (1984) reported cohesion of Sample 3 "minimally related" to combat readiness. Correlations calculated for sample size of lowest level of significance. *Nelson & Berry (1968) reported cohesion "unrelated" to outcomes. *Anomalous result of Pleban et al. (1987) due to small sample size and use of rank order correlation. Groups ranked identically on both cohesion and performance measures.

(31%) of the 39 cases. Another 6 of the samples (15%) focused on the company or battalion level. Among the group levels classified as "other," 18 (46%) were squadrons, detachments, aircrews, tank teams, ship crews, and combinations of levels. Three studies did not report a group level.

Samples. A large majority of the samples (34, or 87%) were active military units. Two samples were reserve units, and 3 samples involved students. Most of the samples (33, or 85%) were from the U.S. military (25 of which were Army units), and 6 samples involved foreign military personnel (5 were Israeli, and 1 was Canadian). Almost all the studies (33, or 92%) included enlisted personnel as participants. Ten samples included officers, and 4 had students as participants.

Cohesion measures. Most studies used items generated by the researchers, sometimes in conjunction with items or scales taken from other measures. However, 16 studies or samples used instruments we judged to be standardized. These measures included the Army Research Institute Platoon Cohesion Index, the Moos instruments (the Military Environment Inventory or the Military Company Environment Inventory), the Israeli Combat Readiness Morale Questionnaire, and the Fundamental Interpersonal Relations Orientation-Behavior.

Researchers reported the reliability of their cohesion measures for 23 of 39 samples (59%). Reliabilities ranged from .57 to .98, with a mean reliability of .78.

Outcome measures. The number of outcome measures per study ranged from one to seven. Of the 39 samples, about half (19, or 53%) used one outcome measure. The modal number of outcomes was 2.0, and the mean was 2.19. For analysis, the outcomes were grouped into the seven categories described previously.

Among the set of outcome variables, eight were standardized measures. These included measures of psychological well-being, physical well-being, and an Army satisfaction scale. In addition, we classified as standardized several Army measures of performance that are used to evaluate individual and group performance, such as Physical Training and Army General Testing and Evaluation Program scores. Reliabilities were reported for 24 outcome measures, with a mean reliability of .76, and a range from .43 to .95.

About half the outcome measures came from the same source as the cohesion measure. Researchers reported using the same source (usually a survey) for collecting both cohesion and outcome data for 46% of the outcomes. The other 54% of the outcomes relied on different sources to obtain these data. Cohesion measures were based on soldier responses, but some of the outcomes came from records data or ratings by commanders or others. Outcome measures for variables such as job satisfaction or well-being (e.g., stress levels or marital satisfaction) were generally self-report measures.

TABLE 2
Summary of Study Characteristics for Individual Samples

<i>Characteristic</i>	<i>Description</i>
Number of samples	Total = 39
Publication dates	Range = 1951–1991 1950s = 4 1960s = 1 1970s = 5 1980s = 25 1990s = 3 In preparation = 1
Publication form	Journal article, book, or test manual = 17 Conference paper, institutional report, or dissertation = 18 Unpublished paper = 4
Number of participants	Total $N = 37,226$ Range = 24–6,748 $M = 954.51$ $Mdn = 547$
Number of groups	Total $N = 1,038$ Range = 4–115 $M = 34.6$ $Mdn = 24.5$
Level of group	Squad, section, or platoon = 12 Company or battalion = 6 Other = 18 Not reported = 3
Sample type	Active component = 34 Reserve component = 2 Students = 3
Country	United States = 33 Israel = 5 Canada = 1
Military service	U.S. Army = 25 Other U.S. military = 6 Foreign = 6 Other = 2
Rank or grade of participants	Enlisted = 36 Officers = 12 Students = 3
Cohesion measures	Mean reliability = .78 (based on 23 samples) Platoon Cohesion Index = 7 Military Environment Inventory or Military Company Environment Inventory = 5 Combat Readiness Morale Questionnaire = 3 Fundamental Interpersonal Relations Orientation–Behavior = 1 Other = 23

(continued)

TABLE 2 (Continued)

<i>Characteristic</i>	<i>Description</i>
Outcome measures	Mean reliability = .76 (based on 24 samples) Mean $N = 2.15$ Range = 1–7
Quality of study rating	$M = 11.02$ $SD = 2.64$ Range = 7.0–16.0
Level of analysis	Individual data, not aggregated = 39 Individual data, aggregated to group level = 21 Group outcome data = 22

Level of analysis of outcomes. We had level-of-analysis data for 82 outcomes. About half (43 of 82, or 52%) of these constituted group-level outcomes or individual data aggregated to the group level. The remaining levels of analysis (39 of 82, or 48%) involved individual data that were correlated without aggregation for the entire sample. Two studies reported data using both approaches.

Interrater reliability. Interrater agreement for 13 nominal variables averaged 89%. The percentage of agreement ranged from 28% to 100%. We obtained 100% agreement for publication date of studies, whereas the 28% rate was for type of sampling plan. The next lowest percentage of agreement for the two coders was 80% for publication form. Without the sampling plan variable, mean interrater agreement for 12 variables was 94%.

For 11 quantitative variables, excluding effect sizes, the correlation between the two raters was .82. The smallest correlation (.67) was for the reliability of the cohesion measure, whereas the correlation for total sample was near 100%. Because total sample was an important variable, agreement on which number to use as the sample was arrived at during conferences throughout the coding process. The correlation for 85 effect sizes was .98. Differences in effect sizes were very small and were often due to rounding differences.

Effect Size Results

Tables 3 and 4 summarize the effect sizes for each outcome class. Table 3 is based on the number of participants, and Table 4 is based on the number of groups. Both tables show the number of cases on which each effect size is based, the total number of participants or groups involved, and the confidence intervals (CI) for the effect size estimates. The tables also contain the unweighted effect size means and the coefficient of robustness (CR; Rosenthal, 1995) for each outcome.

TABLE 3
Effect Sizes and Confidence Intervals Using Number of Participants

<i>Outcome</i>	<i>Number of Cases</i>	<i>Number of Participants</i>	<i>Weighted Effect Sizes^a</i>	<i>Confidence Interval</i>	<i>Unweighted Means</i>	<i>Coefficient of Robustness</i>
Group performance	16	8,152	.400	.378-.422	.428	1.77
Individual performance	9	15,905	.196	.180-.211	.264	1.02
Job/military satisfaction	9	11,361	.470	.451-.488	.451	2.56
Retention	7	14,810	.221	.205-.222	.213	2.42
Well-being ^b	7	9,053	.245	.224-.265	.299	1.44
Readiness	6	3,155	.286	.252-.321	.250	1.34
Indiscipline ^b	4	7,198	.152	.129-.176	.270	1.23

Note. Total $N = 34,555$.

^aWeighted by number of participants. ^bNegative outcomes reverse coded.

TABLE 4
Effect Sizes and Confidence Intervals Using Number of Groups

<i>Outcome</i>	<i>Number of Cases</i>	<i>Number of Groups</i>	<i>Weighted Effect Sizes^a</i>	<i>Confidence Interval</i>	<i>Unweighted Means</i>	<i>Coefficient of Robustness</i>
Group performance	16	577	.331	.250-.412	.428	1.77
Individual performance	7	135	.310	.132-.470	.289	1.15
Job/military satisfaction	7	273	.453	.335-.572	.470	2.47
Retention	4	189	.255	.112-.397	.258	4.37
Well-being ^b	3	145	.334	.174-.497	.450	2.01
Readiness	3	52	.317	.045-.589	.230	1.00
Indiscipline ^b	4	169	.228	.077-.379	.270	1.23

Note. Total $N = 34,555$.

^aWeighted by number of groups. ^bNegative outcomes reverse coded.

In both tables, the number of cases for each effect size ranged from 4 (indiscipline) to 16 (group performance). However, for all but these two outcomes, the number of cases dropped when we weighted by the number of groups (Table 4). The effect size increased in five of seven outcomes and the CI broadened, as a result of weighting by the number of groups. In general, we considered results based on fewer than six or seven cases very questionable, although the pattern of contributing cases is also important in evaluating the result.

CI is used to assess the confidence one can place in a statistic, with a smaller range of values suggesting that greater confidence is warranted. Rosenthal (1995)

considered the CR an additional aid in interpreting the results of meta-analytic reviews. This statistic, which is the mean effect size divided by its standard deviation, becomes larger as variability decreases, the mean effect size increases, or both. Hence, it is possible to use this ratio as an indication of the consistency of the effect size results.

Group performance. This effect size was based on the largest number of cases (16) for both weighting procedures. It was substantial when calculated using the number of participants (.400) but decreased when calculated using the number of groups (.331). The CI was .378–.422 for weighting by participants and .250–.412 for weighting by groups.

Individual performance. The effect size for individual performance suggests a somewhat weaker and less stable relation with cohesion than does the effect size for group performance. However, the effect size of .196 obtained from weighting by participants increased to .310 when weighted by groups, although the CI widened to .132–.470 from .180–.211. The number of cases dropped from nine to seven as a result of the weighting change.

Job/military satisfaction. Despite this drop in the number of cases, this effect size was the largest and most consistent of all the outcomes. The value weighted by participants (.470) decreased somewhat when weighted by groups (.453), and the CI broadened (from .451–.488 to .335–.572).

Retention. Results for this very stable outcome increased slightly as a result of differential weighting (from .221 to .255), although the CI became wider (from .205–.222 to .112–.397). Although the number of cases dropped from seven to four as a result of weighting by the number of groups, the already sizable CR of 2.42 increased to 4.37.

Well-being. The seven original cases for this variable dropped to three when we used weighting by groups, and the effect size rose to .334 from .245. As a result of the weighting change, the CI broadened from .204–.222 to .112–.397.

Readiness. The effect size for readiness was .286 weighted by number of participants and .317 weighted by number of groups. The former value was based on six cases, and the latter based on three cases. The CI became considerably wider, moving from .252–.321 to .045–.589 because of the change in weighting.

Indiscipline. This variable was based on four cases for both types of weighting. As a result of weighting by groups, the effect size rose from .152 to .228, and the width of the CI increased from .129–.176 to .077–.379.

DISCUSSION

Relation of Results to Research Questions

In reviewing our findings, note that the effect size used (r) assesses the degree of a relation and does not imply that one variable necessarily causes another.

Relation of cohesion to outcomes. None of the 95% CIs for the seven mean effect sizes included zero or any minus values, indicating that the relation between cohesion and the various outcomes is a positive one. Because some findings are based on only a few cases, we are more confident about the results for certain outcomes than for others.

We found, for example, that group cohesion was substantially related to soldier perceptions of job and military satisfaction. This robust finding held regardless of whether the mean effect size was weighted by number of participants or number of groups. Because the various cohesion and satisfaction ratings were all self-report measures, use of the same method may have affected these results.

Cohesion was also solidly associated with performance, with group performance more strongly correlated with cohesion than was individual performance. Group performance results also appeared to demonstrate greater consistency than individual performance results as indicated by the CRs for both outcomes.

The mean effect size for indiscipline, which was based on four cases regardless of weighting procedure, was not only based on very few cases but also involved infrequently occurring events. For these reasons, we find the results for indiscipline inconclusive.

The other outcome variables were also based on relatively fewer cases, especially when weighted by groups, and hence our conclusions must be tentative. However, we conclude that group cohesion was positively related to retention, well-being, and readiness. It is interesting to note that, when weighted by groups, the correlation with cohesion increased for these three outcome measures and that the CRs for retention suggested considerable stability for this outcome in particular.

Similarity of results across different types of military units. We observed no clear effect of military service (e.g., U.S. Army, Navy, etc.) or country (United States, Canada, or Israel), but because we had so few samples for such comparisons, our conclusions must be tentative.

Comparability of measures. As shown in Table 1, a wide variety of both cohesion and outcome measures has been used in military research. Some of the cohesion measures (especially the sociometric ones) focus on the interpersonal aspect of the phenomenon, but others also incorporate the task or teamwork dimension of cohesion. We attempted to code the studies for different dimensions but abandoned the task because few researchers provided this kind of information. Sometimes, they reported using a particular instrument, but our examination of the data revealed that they used only some items from the measure. Thus, we found it difficult to make comparisons of results supposedly involving the same instrument.

Future research may shed further light on the comparability issue. For now, we suggest that a varied set of measures appears to result in similarly inconclusive findings.

Relation of results to previous meta-analyses. Our research resulted in effect sizes roughly comparable to those obtained in other meta-analyses. For example, our group performance results were close to the results of Evans and Dion (1991), whose r values were .36 uncorrected and .42 corrected.

Gully et al. (1995) reported a range of effect sizes (.166 to .464) that varied according to whether they were corrected for unreliability of measurement, involved self-report measures of performance, or took into account the moderators of level of analysis³ or degree of task interdependence. Although our mean effect sizes tended to be lower than those reported by Gully et al., our values for group performance (.400 weighted by participants, .331 weighted by groups, and .428 unweighted) were fairly similar to their results for high levels of task interdependence (.387 uncorrected and .464 corrected). Although we did not code degree of task interdependence, we did separate performance on the basis of group and individual tasks. Thus, group tasks, such as platoon performance on a field problem, would be more interdependent than individual tasks, such as scores on a job knowledge task or rifle qualification exercise.

Our effect sizes for performance were higher than those Mullen and Copper (1994) obtained for their set of military groups. In addition to rejecting 44 of the 10 studies Mullen and Copper classified as military, we differentiated between group and individual performance, which Mullen and Copper did not. When we averaged the effect sizes for individual performance and group performance, we obtained an overall mean effect size for performance of .265 when weighted by number of partic-

³We did not have enough cases for most outcomes to compare group- and individual-level analyses as Gully et al. (1995) did. However, we obtained an effect size of .439 for the 12 cases of group performance that used the group as the unit of analysis. The 4 cases of group performance that correlated individual cohesion measures directly with individual performance measures resulted in an effect size of .347.

⁴Of these 4 studies, 2 included civilian as well as military personnel, 1 did not involve an ongoing group, and 1 lacked a measure that met our definition of cohesion.

ipants and .325 when weighted by number of groups, bringing our results closer to those of Mullen and Copper, who obtained .23 for their "military" subgroup.

Because previous meta-analyses have considered only performance as an outcome measure, we cannot compare our results for the other outcomes.

Implications of Results for Military Policy and Planning

In general, military planners and policymakers believe that cohesion is a desirable characteristic of military units and try to foster its development. Our findings provide substantial empirical support for such activities in terms of performance and job satisfaction outcomes. The relation of cohesion to other outcomes is not yet firmly established, but the available research suggests that cohesion enhances well-being, increases retention and readiness, and works against indiscipline in military groups. We encourage the services to continue their efforts to enhance cohesion and to consider developing (and documenting the effect of) interventions designed to increase cohesion.

For example, to enhance cohesion, some authorities (e.g., Blaufarb, 1989) have argued in favor of lengthening tours to diminish turbulence. Stabilizing personnel supposedly leads to increased cohesion and, thus, to greater combat effectiveness (see Griffith, 1986), although research by Oliver (1988b) suggests that intermediate levels of turbulence may be most desirable in enhancing cohesion and performance, noting that one of the problems with the COHORT (Cohesion, Operational Readiness, and Training) system was that the junior enlisted personnel were well stabilized but their officer leadership was not. When two values such as these—increasing personnel stability and enhancing officer career development—conflict, Army managers must make difficult decisions as to how to resolve these issues.

Implication of Results for Future Research

Our results have many implications for future research, some of which are briefly noted here.

Reporting basic data. Meta-analysts are invariably amazed by how poorly research is reported when they face the task of coding a set of studies (Orwin, 1994). To produce accurate studies, researchers should report means, standard deviations, number of participants, and intercorrelations of all continuous variables. All results need to be included, not just significant ones, with exact probabilities of results reported. Researchers should describe their samples, instruments, and settings in sufficient detail to enable readers to judge the extent to which the research might generalize to other situations and to enable meta-analysts to code pertinent variables. Cohesion researchers also need to report both the total number of participants and the number of groups as well as nonresponse rates and attrition data.

Conceptualization. We felt that the conceptualization in the studies we reviewed tended to be weak. Many researchers did not define what they meant by *cohesion*. They may have quoted classic authorities such as Lott and Lott (1965) but then neglected to specify their own definition. We feel it is essential for researchers to define such terms so that the operationalization of the variables can be assessed.

Selecting measures. We believe that the military cohesion research contains good examples of instruments that have demonstrated their worth and would be useful in future research. There are several cohesion measures, for example, and a number of outcome measures that provide reliability and validity data. We urge researchers to incorporate such measures into their future investigations.

In addition, we recommend that researchers tap different sources for their cohesion and outcome measures. Measures of satisfaction and some of the standardized well-being scales will necessarily be from the same source as the cohesion measures; but in other cases, especially when collecting performance data, researchers should use different sources for their measures.

Level of analysis. We found that cohesion researchers often ignored the level of analysis issue. Correlating individual data with outcomes across the entire sample ignores the group aspect of cohesion. The appropriate level will depend on the researcher's purpose and the mission of the units involved. In general, we recommend lower levels, such as squad or platoon, because larger units may get beyond the "small-group" aspect of cohesion and are less likely to meet definitions of the construct. Sometimes, of course, performance data are available at a higher level—the company road-march scores of Siebold and Kelly (1987), for example, or the ship performance data used by Allen (1986). In both these cases, the researchers used performance data for a higher organizational level that functioned as a group to accomplish a specific mission.

Both cohesion and outcomes need to represent the same level. For example, if the level chosen is the platoon, the cohesion instrument should query respondents about interactions at the platoon level. If the outcome measure is not a measure of platoon performance as a group, individual performance data should be aggregated to the group (platoon) level. Some authorities (e.g., Gully et al., 1995) also emphasize the need to examine the agreement of responses at the individual level before aggregating to a group level. According to Gully et al., aggregation bias becomes a "potentially severe problem" (p. 513) if the homogeneity of individual-level responses is not ensured.

Other moderator variables. Although we coded potential moderators, too few studies used the same variables to make such analyses meaningful. We encourage cohesion researchers to include measures of leadership style, demographic

characteristics, task interdependence, and other potential moderators to enable future meta-analysts to explore their relation to cohesion.

CONCLUSIONS

We conclude that the results of our meta-analysis demonstrate that group cohesion in military units is related in a positive manner to various desirable outcomes of interest to the military services. This meta-analysis is the first such effort limited to the cohesion of military groups and also the first to examine the relation of cohesion to outcomes other than performance. Our results demonstrate that looking at measures of individual performance separately from measures of group performance has merit because these variables are differentially related to cohesion. In addition, we have reported effect sizes weighted by number of groups as well as by number of participants. We also are convinced of the need to consider level of analysis in selecting or developing measures and in conducting data analysis. As noted previously, our findings have relevance both for military planning and for future research.

ACKNOWLEDGMENTS

We thank Stan Gully, Hannah Rothstein, and Betsy Becker for their comments on earlier versions of this article. The views expressed in this article are those of the authors and do not necessarily reflect the views of the U.S. Army Research Institute or the Department of the Army.

REFERENCES

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

- *Allen, N. H. (1986). In search of a "mix" for excellence: An exploratory study of the possible relationship between organizational performance and the psychological type of key group members. *Dissertation Abstracts International*, 47(10), 3808A. (University Microfilms No. 8700545)
- *Allen, N. H., & Potter, E. H. (n.d.). *Identifying high performing systems in the U. S. Coast Guard*. Unpublished manuscript.
- *Blades, J. W. (1976). The influence of intelligence, task ability and motivation on group performance. *Dissertation Abstracts International*, 37(03), 1463A. (University Microfilms No. 76-20, 706)
- *Blades, J. W. (1986). *Rules for leadership: Improving unit performance*. Washington, DC: National Defense University Press.
- Blaufarb, P. N. (1989). *Air Force officer cohesion* (Air War College Research Report). Maxwell Air Force Base, AL: U.S. Air Force, Air University. (NTIS No. AD-A217 655)
- Carron, A. V. (1982). Cohesiveness in sport groups: Interpretations and considerations. *Journal of Sport Psychology*, 4, 123-138.
- *Cohen, E. (1957). The effect of members' use of a formal group as a reference group upon group effectiveness. *Journal of Social Psychology*, 46, 307-309.

- Datel, W. E. (1978). *Final evaluation report on Fort Ord Project Volar*. Alexandria, VA: Defense Technical Information Center. (DTIC No. AD-A054 006)
- *Dickenson, R. A. (1991). Cohesion in military college recruits. In *Proceedings of the 33rd Annual Conference of the Military Testing Association* (pp. 73-78). San Antonio, TX: U.S. Air Force Armstrong Laboratory Human Resources Directorate and the U.S. Air Force Occupational Measurement Squadron.
- *Dobbins, G. H., & Zaccaro, S. J. (1986). The effects of group cohesion and leader behavior on subordinate satisfaction. *Group & Organization Studies*, 11, 203-219.
- *Downey, R., Duffy, P., & Shiflett, S. (1975). *Criterion performance measure of leadership and unit effectiveness in small combat units* (ARI Research Memorandum No. 75). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences.
- Dupuy, H. P. (1978, October). *Self-representations of general psychological well-being of American adults*. Revision of a paper presented at the meeting of the American Public Health Association, Los Angeles.
- Evans, C. R., & Dion, K. L. (1991). Group cohesion and performance: A meta-analysis. *Small Group Research*, 22, 175-186.
- Fiedler, F. E. (1967). *A theory of leadership effectiveness*. New York: McGraw-Hill.
- *Gal, R., & Manning, F. J. (1984, August). *Correlates of unit cohesion and morale in the U.S. and Israeli Armies*. Paper presented at the 92nd Annual Meeting of the American Psychological Association, Toronto, Canada.
- *Gal, R., & Manning, F. J. (1987). Morale and its components: A cross-national comparison. *Journal of Applied Social Psychology*, 17, 369-391.
- *Goodacre, D. M., Jr. (1951). The use of a sociometric test as a predictor of combat unit effectiveness. *Sociometry*, 14, 148-152.
- *Goodacre, D. M., Jr. (1953). Group characteristics of good and poor performing combat units. *Sociometry*, 16, 113-127.
- *Griffith, J. (1986). *The measurement of military unit cohesion: An empirical approach*. Unpublished manuscript.
- *Griffith, J. (1987, April). *Group cohesion, training performance, social support and the Army's new unit replacement system*. Paper presented at the annual meeting of the Eastern Psychological Association, Arlington, VA.
- *Griffith, J. (1988). Measurement of group cohesion in U.S. Army units. *Basic and Applied Social Psychology*, 9, 149-171.
- Gully, S. M., Whitney, D. J., & Devine, D. J. (1995). A meta-analysis of cohesion and performance: Effects of level of analysis and task interdependence. *Small Group Research*, 26, 497-520.
- *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.
- Hogg, M. A. (1992). *The social psychology of group cohesiveness: From attraction to social identity*. New York: New York University Press.
- Hoiberg, A., & Berry, N. H. (1978). Expectations and perceptions of Navy life. *Organizational Behavior and Human Performance*, 21, 130-145.
- Hoiberg, A., & Pugh, W. M. (1978). Predicting Navy effectiveness: Expectations, motivation, personality, aptitude, and background variables. *Personnel Psychology*, 31, 841-852.
- *Hom, P. W., & Hulin, C. L. (1981). A competitive test of the prediction of reenlistment by several models. *Journal of Applied Psychology*, 66, 23-39.
- Hunter, J. E., & Schmidt, F. L. (1990). *Methods of meta-analysis*. Newbury Park, CA: Sage.
- *Julien, T. D., & Siebold, G. L. (1989). *Unit leader assessment of a joint readiness training center rotation* (ARI Research Rep. No. 1564). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences.
- Lott, A., & Lott, B. (1965). Group cohesiveness as interpersonal attraction: A review of relationships with antecedent and consequent variables. *Psychological Bulletin*, 64, 259-309.

- *Manning, F. J., & Fullerton, T. D. (1988). Health and well-being in highly cohesive units of the U.S. Army. *Journal of Applied Social Psychology*, 18, 503-519.
- *Manning, F. J., & Ingraham, L. H. (1987). An investigation into the value of unit cohesion in peacetime. In G. Belenky (Ed.), *Contemporary studies in combat psychiatry* (pp. 47-67). New York: Greenwood.
- Matt, G. E., & Cook, T. D. (1994). Threats to the validity of research synthesis. In H. Cooper & L. V. Hedges (Eds.), *The handbook of research synthesis* (pp. 503-520). New York: Russell Sage Foundation.
- Moos, R. H. (1973). *Military Company Environment Inventory manual*. Palo Alto, CA: Stanford University, Social Ecology Laboratory.
- *Moos, R. H. (1986). *Military Environment Inventory manual*. Palo Alto, CA: Consulting Psychologists Press.
- *Motowildo, S. J., & Borman, W. C. (1977). Behaviorally anchored scales for measuring morale in military units. *Journal of Applied Psychology*, 62, 177-183.
- Mullen, B., & Copper, C. (1994). The relation between group cohesiveness and performance: An integration. *Psychological Bulletin*, 115, 210-227.
- Neff, J., & Husaini, B. (1980). Race, socioeconomic status, and psychiatric impairment: A research note. *Journal of Community Psychology*, 8, 16-19.
- *Nelson, P. D., & Berry, N. (1968). Cohesion in Marine recruit platoons. *Journal of Psychology*, 68, 63-71.
- Oliver, L. W. (1988a). *The relationship of group cohesion to group performance: A research integration attempt* (ARI Tech. Rep. No. 807). Alexandria, VA: U.S. Army Research Institute.
- Oliver, L. W. (1988b). *The relationship of leader turbulence to platoon cohesion and performance*. Unpublished manuscript. Alexandria, VA: U.S. Army Research Institute.
- Oliver, L. W. (1990). *Cohesion research: Conceptual and methodological issues* (ARI Research Note No. 90-133). Alexandria, VA: U.S. Army Research Institute.
- *Olson, D. M., & Borman, W. C. (1987). *Developmental and field tests of Army Work Environment Questionnaire* (ARI Tech. Rep. No. 737). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences.
- Orwin, R. G. (1994). Evaluating coding decisions. In H. M. Cooper & L. Hedges (Eds.), *The handbook of research synthesis* (pp. 139-162). New York: Russell Sage Foundation.
- Orwin, R. G., & Cordray, D. S. (1985). Effects of deficient reporting on meta-analysis: A conceptual framework and reanalysis. *Psychological Bulletin*, 97, 134-147.
- *Pleban, R., Valentine, P. J., & Thompson, T. (1987). *Enhancing small group cohesion and effectiveness in long range reconnaissance teams* (ARI Research Note No. 30). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Science.
- Procidano, M. E., & Heller, K. (1983). Measures of perceived social support from friends and from family: Three validation studies. *American Journal of Community Psychology*, 11, 1-24.
- Roach, A. J., Frazier, L. P., & Bowden, S. R. (1981). The marital satisfaction scale: Development of a measure for intervention research. *Journal of Marriage and the Family*, 43, 537-546.
- Rosenthal, R. (1994). Parametric measures of effect size. In H. M. Cooper & L. Hedges (Eds.), *The handbook of research synthesis* (pp. 139-162). New York: Russell Sage Foundation.
- Rosenthal, R. (1995). Writing meta-analytic reviews. *Psychological Bulletin*, 118, 183-192.
- *Royle, M. H. (1985). *Factors affecting attrition among Marine Corps women* (Tech. Rep. No. 86-7). San Diego, CA: Navy Personnel Research and Development Center.
- Shadish, W. R., & Haddock, C. K. (1994). Combining estimates of effect size. In H. Cooper & L. V. Hedges (Eds.), *The handbook of research synthesis* (pp. 261-281). New York: Russell Sage Foundation.
- *Shiron, A. (1976). On some correlates of combat performance. *Administrative Science Quarterly*, 21, 419-432.
- Shutz, W. C. (1958). *FIRO: A three-dimensional theory of interpersonal behavior*. New York: Rinehart.
- *Siebold, G. L. (1990, November). *Cohesion in context*. Paper presented at the 33rd Annual Conference of the Military Testing Association, Orange Beach, AL.

- *Siebold, G. L., & Kelly, D. R. (1987, October). *The impact of unit cohesion on unit performance, morale, and ability to withstand stress: A field exercise example* (Working Paper No. LM 87-13). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences.
- *Siebold, G. L., & Kelly, D. R. (1988a). *Development of the Platoon Cohesion Index* (Tech. Rep. No. 816). Alexandria, VA: U.S. Army Research Institute for Behavioral and Social Sciences.
- *Siebold, G. L., & Kelly, D. R. (1988b). *The impact of cohesion on platoon performance at the Joint Readiness Training Center* (Tech. Rep. No. 812). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences.
- *Solomon, Z., Mikulincer, M., & Hobfoll, S. E. (1986). Effects of social support and battle intensity on loneliness and breakdown during combat. *Journal of Personality and Social Psychology*, 51, 1269-1276.
- *Sterling, B., & Williams, L. (1982, June). *A study of cohesion in Army at units at four unit levels* (ARI working paper). Fort Benjamin Harrison, IN: Field Unit, U.S. Army Research Institute for the Behavioral and Social Sciences.
- Taylor, J. C., & Bowers, D. G. (1972). *Survey of organizations: A machine scored standardized questionnaire instrument*. Ann Arbor: University of Michigan, Institute of Social Research.
- *Turney, J. R., & Cohen, S. L. (1976). *The development of a Work Environment Questionnaire for the identification of organizational problem areas in specific Army work settings* (ARI Tech. Rep. No. 275). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences.
- *Tziner, A., & Vardi, Y. (1983). Ability as a moderator between cohesiveness and tank crews performance. *Journal of Occupational Behaviour*, 4, 137-143.
- *Vickers, R. R., Wallick, M. T., & Hervig, L. K. (1982). *The Marine Corps basic training experience: Stresses, leadership, and group cohesion as predictors of attitudes, health, and performance* (Rep. No. 82-28). San Diego, CA: Naval Health Research Center.
- Ware, J. E., Johnson, S., Davies-Avery, A., & Brook, R. (1979). *Mental health: Vol. 3. Conceptualization and measurement of health in adults in the Health Insurance Study*. Santa Monica, CA: Rand.
- *Yagil, D. (1995). *A study of cohesion and other factors of major influence on soldiers' and unit effectiveness* (ARI Research Note No. 95-11). Alexandria, VA: U.S. Army Research Institute.
- *Yost, E. E., & Tremble, T. R., Jr. (1984, August). *Impact of cohesion on leader behavior-outcome relationships*. Paper presented at the 92nd Annual Meeting of the American Psychological Association, Toronto, Canada.
- *Zaccaro, S. J. (1991). Nonequivalent association between forms of cohesiveness and group-related outcomes: Evidence for multidimensionality. *Journal of Social Psychology*, 13, 387-399.