

C

E



Center for
Effective
Organizations

**TESTING FOR GROUPNESS:
A THEORY- BASED APPROACH TO
AGGREGATION ISSUES IN WORK
GROUP RESEARCH**

**CEO PUBLICATION
T 00-2 (375)**

DIANE E. BAILEY
Stanford University

EILEEN M. VAN AKEN
Virginia Tech

SUSAN G. COHEN
University of Southern California

March, 2000

**Testing for Groupness: A Theory-Based
Approach to Aggregation Issues in Work Group Research**

Diane E. Bailey*
Department of Industrial Engineering and Engineering Management
Stanford University

Eileen M. Van Aken
Department of Industrial and Systems Engineering
Virginia Tech

Susan G. Cohen
Center for Effective Organizations
University of Southern California

* Corresponding author. Send inquiries to Diane E. Bailey, Department of Industrial Engineering and Engineering Management, 490 Terman, Stanford University, Stanford, CA 94305-4024, email Diane.Bailey@stanford.edu, phone 650(723)-3821.

An earlier version of this article was presented at the Industrial Engineering Research Conference, Vancouver, Canada, May 1997.

Data used in this study were collected from previous research conducted by Susan G. Cohen and Allan M. Mohrman. We are grateful to Alice Yee Mark and Beth Neilson for their technical data support.

Abstract

In this paper, we address the issue of aggregating individual employee level data to the work group level. In contrast to recent literature that focuses on statistical justification of various aggregation methods, we take a theoretical perspective centered on key variables used in the definition of work groups. We argue that for these key variables, individual responses should exhibit low within-group variation and high overall values if the unit is to be considered a work group. We illustrate our argument with data from the research division of a petrochemical corporation. This work adds to the current debate of proper methodology in team research, and has practical implications for managers seeking to establish or define teams.

115 words.

Testing for Groupness: A Theory-Based Approach to Aggregation Issues in Work Group Research

Recently, a methodological and theoretical debate has arisen in the work group literature. The heart of the controversy centers on the proper method to use when aggregating individual-level responses to the group level. Such aggregation is common in studies of work groups. For example, a researcher might distribute a survey to be completed individually by employees. The responses of all members of a given work group are then aggregated together (typically by simple averaging) to achieve a single group value. Only rarely is a single survey given to a group as a whole, to be completed by the group acting in consort.¹ Often the averaging of individual responses is performed without regard to variation among them, but more recently greater attention has been paid to evaluating the appropriateness of aggregating disparate individual responses.

One point of contention in the current debate is whether James, Demaree, and Wolf's (1984) interrater reliability coefficient is sufficient to determine the appropriateness of aggregation. Yammarino and Markham (1992) contend that it is not; rather, they argue, within-and-between-analysis (WABA) should be employed instead (e.g., as in Van Aken & Kleiner, 1997, to examine determinants of cross-functional team effectiveness). Other researchers pose further statistical arguments against the use of James et al.'s measure (Schmidt & Hunter, 1996, 1989). Klein, Dansereau, and Hall (1994), in a paper about levels of analysis, raise the argument

¹ See Gibson, Randel, and Early (in press) and Kirkman, Tesluk, and Rosen (1998) for recent comparisons of the predictive abilities of group-completed surveys versus aggregated individual surveys.

from a debate over proper statistical methods to one of proper theory. Rousseau (1985) also argues that theory should guide the choice of level of analysis in organizational research.

It is exactly an issue of theory that we wish to take up in regard to this debate. What we question is not the statistical measure that should be employed in aggregating work group responses. Rather, we wish to question the idea that if one finds—by whatever method—low within-group variation (i.e., agreement) in regard to a certain variable(s), then the individuals can be considered a work group. It seems to us that one should have at the outset some concept of what a work group is, and then, with that concept in hand, check the organizational unit under study to confirm that its features match this concept. That is, one should identify certain variables that capture the essence of what it is to be a work group, then look for low within-group variation, and sufficiently high values, in individual responses with regard to these variables. For example, if one deems task interdependence to be central to one's concept of what it means for a group of employees to be a work group, then the responses to questions of task interdependence of the individuals in the unit under study should be rather consistent (i.e., they should exhibit low within-group variation). They also should be sufficiently high in value (in this case, they should indicate that the individuals are quite dependent on one another in completing their tasks). Such tests of the data constitute validity checks in the sense that they provide confirmation that the units under study are true work groups.

Confirming that work groups exist is quite different from confirming that an individual construct exists at the group level, which is what the previous authors attempted to do and what gave rise to their methodological disagreements. We certainly are not challenging the existence of any number of variables at the group level. For example, a group of employees jointly may view their environment in the same manner, such that shared perception of the work climate

becomes a construct with meaning for the group. Similarly, one can find consensus among the congregation of a church in the basic doctrines of the religion espoused there. Our point is that just because a group of people at work have shared ideas about their climate no more makes them a work group than having a shared belief makes a church congregation one. Thus, confirming that the units under study are work groups is the first requirement for any study of work groups, to be followed by confirmation of the particular variables of concern (e.g., shared climate, conflict, participation) in relation to the theory and hypotheses of the research.

This paper is organized in the following way. First, we define what we mean by a work group. Second, we provide an example from the group literature to illustrate the importance of confirming that the units studied are work groups. Then, drawing upon the definition of a work group, we highlight key variables—such as task interdependence—that should have low variance among individual responses and a sufficiently high level if the unit under study is to be considered a work group. We discuss how other group-related variables associated with work group effectiveness (e.g., conflict or participation) might have large individual variation even within a work group without detracting from the unit's “groupness.” Finally, using survey data collected from employees of a large petrochemical corporation, we demonstrate a procedure for confirming work group status. This work adds to the discussion of proper methodology in group research, and has practical implications for managers seeking to establish or define teams.

Defining a Work Group

Exactly what do we have in mind when we use the term “work group”? Perhaps the most accepted definition of a work group in the literature comes from Hackman (1987): a work group is a collection of individuals who are interdependent in their tasks, who share responsibility for outcomes, who see themselves and who are seen by others as an intact social entity embedded in

one or more larger social systems, and who collectively manage their relationships across organizational boundaries.

Hackman's definition builds upon earlier understandings of groups. For example, Lewin (1948) defines a group as a dynamic whole based on interdependence rather than similarity. Cartwright and Zander (1968: 46) define a group as a "collection of individuals who have relations with one another that make them interdependent to some significant degree." Merton (1957) defines a group as a number of people who interact with one another in accord with established patterns. He adds that the interacting persons must define themselves as members of the group, and that they and others must differentiate between group members and non-members. Hackman's concepts of work group interdependence and being viewed as an intact social unit are derived from this earlier work. Alderfer (1977), taking an open systems perspective, adds a further criterion to the definition of a group: members acting alone or in concert must have significant interdependent relations with other groups. Hackman adapts this criterion by contending that work group members collectively must manage their relationships across organizational boundaries.

What Hackman adds to these earlier definitions is the focus on work groups, or groups that exist within an organizational context. Work groups have one or more tasks to perform; these tasks lead to identifiable and potentially measurable group outcomes. Work group members may belong to one or perhaps several organizations, but they always operate within an organizational context. Researchers who have followed Hackman have engaged similar definitions of work groups (e.g., Guzzo and Dickson, 1996; Sundstrom, DeMeuse, and Futrell, 1990; Cohen and Bailey, 1997).

Its derivation from a long history of group research and its acceptance among current group researchers make Hackman's definition a natural choice as the basis for any effort to confirm the work group status of organizational units under study. Consistent with recent research (Chan, 1998; Cohen and Bailey, 1997), we do not differentiate between the terms “group” and “team” in this paper. Hackman's definition and the procedure we build based upon it are meant to apply to both terms. Specifically, Hackman's definition applies to traditional work teams, which are continuing work units responsible for producing goods or providing services. It also applies to parallel teams (e.g., quality circles and process improvement teams), project teams (e.g., cross-functional product development teams), and management teams. We expect this definition to hold true even for newer organizational work forms, such as virtual teams whose memberships span time, space, organization, or combinations thereof. For a traditionally collocated team, the larger social system may be the organization or even the shift on which the team works. For a virtual team, the larger social system may be the set of suppliers, customers, and manufacturers that constitute a supply chain. Both teams will be recognized as intact social entities within their respective social systems, it is just that the social system for the virtual team is broader in scope and scale than that for the traditional team. In sum, Hackman's definition of a work group applies equally well across team types, notwithstanding the differing interpretations of its elements according to the type of team examined.

Of course, some variation is to be expected among teams, and perhaps even among types of teams, in the degree to which they fulfill each aspect of Hackman's definition. For example, we may suspect that certain management teams constitute work groups more in name than in reality. We could determine the truth of such a suspicion by judging the extent that the managers are independent in their tasks rather than dependent, that they individually rather than

collectively manage their boundaries, and so forth. For this reason, "groupness" is not a binary variable but rather a continuous one. Some work units on the far end of the continuum will clearly meet the elements of the definition put forth; no one would challenge their groupness. Others at the opposite end will just as clearly fail in meeting the definitional requirements, such that no one would claim work group status for them. Other groups may fall somewhere in the middle range. In such cases, qualitative descriptions of the work tasks and work environment will help readers of the research to determine to what extent the study of these units has relevance to theories and models of work groups.

Why Work Group Confirmation is Needed

One question that is asked of every study of work groups is how generalizable are its findings to other work groups. Generalizability is a major concern because scholars build theories and models of work group effectiveness on the presumption that they will apply more or less equally well to all work groups. In the absence of generalizability, an idiosyncratic model must be built for every work group studied, and universal theories escape us. A precondition for generalizability is that the study's sample actually consists of work groups. While such a precondition may seem obvious, several recent studies have appeared in which the work group status of the units under study is highly questionable. We discuss one such study here as an example of the types of problems it poses for theory and model-building.

Yammarino and Dubinsky (1990) investigated the relationships between three categories of managerially controllable factors (supervisor behaviors, autonomy, and role-related factors) and various measures of salesperson performance. They conducted their analysis at both the individual and the group level with a sample drawn from two industries: retail sales and insurance. In testing 36 different relationships between individual managerially controllable

factors and distinct outcomes across the two samples, the authors found 20 relationships to be significant at the individual level but not the group level, and 10 relationships to be significant at neither level. Only six relationships were significant at the group level of analysis. On the basis of these findings, the authors recommended that the employees should be treated as individuals in terms of supervisory behaviors and autonomy, but that they should be treated and managed as groups in terms of role ambiguity and role problems.

Yammarino and Dubinsky framed their study as an investigation of work groups, noting that prior studies of salesperson performance in relation to managerially controllable factors focused only on the individual as the unit of analysis. Concerns of generalizability take on special importance for this work because one of their findings in particular, namely that autonomy was not related to performance ratings for the insurance groups, stands in contrast to a number of recent studies in which group autonomy was found to be positively related to performance for work groups (e.g., Cohen & Ledford, 1994; Campion et al., 1993; Pearson, 1992). More interestingly, it also stands in contrast to their own results for retail sales groups, for which autonomy and performance were positively related.

The contradictory evidence provided by Yammarino and Dubinsky's study is intriguing, but what is troubling in this case is that the groups of salespersons don't appear to be work groups. The groups were collections of individuals who seemed to lack any interdependence or collective responsibility, two major elements in the definition of work groups. The authors defined a work group as the union of all salespersons who reported to the same manager. No further check for the status of these collections as work groups was made. Other studies (e.g., George, 1990; George and Bettenhausen, 1990; and George, 1995) have constructed similar samples on similar premises; namely, they have construed *a priori* work groups based on

configurations of retail salespersons who report to the same manager. We hesitate to draw conclusions for work groups on studies that claim to study work groups, but whose samples and methodology fail to convince us that they do. The implicit dangers for theory development and managerial practice are quite clear.

We point out problems in the work by Yammarino and Dubinsky and George and her colleagues because these studies and their authors have been at the center of the debate over how best to statistically justify aggregation of individual-level data to the group level. We commend these authors for their detailed and conscientious efforts to justify aggregation of the variables that were of particular concern to them, such as affective tone in George (1990) and managerially controllable factors in Yammarino and Dubinsky (1990). Their work in this regard provides valuable direction for the field. Their failure to extend these efforts to variables that might convince us that their samples consisted of work groups, as opposed to collections of individuals at work who agree about certain items, is what concerns us.

Certainly, many other studies omit statistical justification of the existence of true work groups in their samples. Quite a number of these studies, however, do provide very thorough qualitative descriptions of the work groups that allay any qualms over the nature of the sample. [See Batt and Appelbaum (1995) for an excellent example.] We strongly encourage all group researchers to provide such descriptions. In cases where the qualitative evidence is missing or insufficient, however, there are quantitative procedures that can be employed to help verify that the units under study are true work groups. We intend here to explain one such procedure that is based upon Hackman's definition of a work group, and to demonstrate its use.

In this regard, our paper is concerned primarily with research that incorporates surveys as part of its design. We don't argue that all studies of work groups should include surveys and

follow testing procedures like the ones we propose here, but rather that they be sensitive to the kinds of definitional elements we discuss. For example, qualitative studies might provide evidence that their units are work groups via analysis of organizational artifacts, quotes from interviews, or descriptions of member behavior in team meetings. In essence, the argument we make—that studies of work groups should provide evidence that the collections of employees in their sample are in fact work groups—is applicable to all group researchers, whereas the techniques we suggest are aimed more specifically towards survey research.

Testing for “Groupness”

A researcher has several means available to test whether or not an organizational unit under study is in fact a work group. For example, the researcher can conduct interviews in which individuals are asked to identify the members of their work group, to discuss the nature of their tasks, and to specify who is responsible for maintaining various organizational relationships. Consolidation of the interviews, with follow-up checks in the case of discrepancies, should yield the relevant work groups at a site. Triangulation of data sources also may prove useful; managers, co-workers, outside contractors, customers and the like may provide external verification of work group status.

Unfortunately, this interviewing process is time-consuming, and would no doubt prove infeasible for researchers employing survey research methods over a large sample. Even these researchers, however, can verify that their units are work groups by testing for groupness among the individual responses they receive. In accordance with Hackman's definition of a work group, four key elements should be examined: task interdependence, shared responsibility, sense of being an intact social entity, and collective boundary management. For the unit under study to

qualify as a work group, there should be low within-group variation and sufficient values for these four variables.

Task interdependence. The first part of the definition states that a work group is a collection of individuals who are interdependent in their tasks. Task interdependence requires people to work together to get the job done. Thus, task interdependence is a key variable that should be measured and compared within groups, and often is. If responses suggest that task interdependence is very low within a unit, then, even if the individual responses reflect agreement on this point, the researcher must question the unit's groupness. Furthermore, if there is great variation in the responses, then the researcher should examine the individual responses. It may be that a core work group exists, but that peripheral members have been included who are not truly part of the group. Task interdependence is, in our opinion, the key variable for defining a work group. Any problems in consistency and value in regard to this variable should raise definitional concerns for the researcher.

Shared responsibility. The second part of the definition specifies that a work group's members share responsibility for outcomes. How can shared responsibility be determined? Rarely in studies of groups have we seen variables that explicitly capture shared responsibility, which is a bit surprising. The Campbell-Hallam Team Development Survey (Hallam & Campbell, 1994) assesses a related construct called "team commitment" that is measured by capturing, among other things, the extent to which members are committed to superior team performance and accept personal responsibility for the success of the team. Seemingly, a survey could include questions that directly ask the extent to which members of the respondent's work group share responsibility for specific outcomes (e.g., productivity and quality goals for traditional work teams, time-to-market for new product development teams). Certain proxy

variables also may capture the essence of shared responsibility, although they are less preferred, being more distant from the construct. Mission clarity and goal acceptance are two possible proxy variables, as they suggest that the group understands and accepts its collective purpose. Self-perceptions of group performance are another possibility, as they reflect the group's awareness of its outcomes. As with interdependence, researchers should be concerned not only that there is low variance among individual responses to questions concerning shared responsibility, but also that the reported levels of shared responsibility are sufficiently high.

Intact social entity. The third part of the definition suggests that work group members must see themselves and be seen by others as an intact social entity embedded in one or more larger social systems. The latter part of this definition need not be verified in most cases, as typically the units under study have been drawn from an organization(s), which constitutes the larger social system. The part that must be verified is that the units are intact social entities.

Here, the example of individual interviews in which respondents identify their team members seems exceptionally appropriate. Several alternatives that lend themselves to survey methods may also work. Perhaps the most obvious route, but one more tedious for analysis, is to include a survey question that simply asks the respondent to list all members of the work group, including the respondent. Alternatively, the survey might ask the respondent how many members are in the work group, or simply the name of the group. These approaches require the respondent to have a sense of what is meant by a work group. Given that many firms now employ a wide variety of teams and groups, often with idiosyncratic definitions, these questions should be preceded with a description of what is meant by a work group, which in any event seems a good idea for a survey on work groups. If responses to questions about membership, size, and/or group name are quite disparate, then it would seem that the unit under question lacks

a clear vision of itself as an intact social entity. Previous research has utilized related concepts such as “group identity;” however, these studies seem to fall more within the realm of social groups rather than work groups (for example, Murrell & Dietz, 1992, discuss common group identity in fans of sports teams).

Collective boundary management. The fourth part of the definition is perhaps the most difficult to test; it specifies that a work group must collectively manage its relationships across organizational boundaries. What does it mean to manage one's relationships, and to do so across organizational boundaries? In the case of a production work group, this might mean that agreements on equipment sharing with other groups cannot be made by an individual on behalf of the group without the group's consent. Together, the group determines how it will share equipment with other organizational entities or it authorizes one of its members to act on its behalf.

If units within the sample are similar in regard to the nature of their work, then a researcher may be able to determine, through initial interviews and observation at a pilot site, the predominant relationships that a typical work group possesses. A subsequent survey question could list those relationships and ask respondents to indicate to what extent the group manages them. In the absence of detailed information of this sort, or in cases where the nature of the work varies considerably across the sample, a more general question may be appropriate (e.g., "To what extent does your work group manage its external relations (customers, suppliers, supervisors, or other work units within the firm) as a group as opposed to individually?") Finally, several proxy variables, such as external communication, conflict, or cooperation, may serve as an indication of the groups' collective managing of external relationships. Once again, researchers should be concerned not only with low variance among individual responses to

questions about boundary management, but also with the level of boundary management that is reported. If the level is low, then the unit's degree of groupness should be questioned.

Interpretation of these four definitional elements. We wish to emphasize that in order for a collection of individuals within an organization to be considered a work group, low within-group variation ideally should be found among individual responses to questions for all four elements described above, and the level of response (for example, how much task interdependence) should also be examined. An organizational unit should not be considered a work group just because one finds low variation among its members' responses to other variables related to the group's design, processes, or outcomes. In such cases, a group level of analysis may be justified in accordance with whatever theory one is investigating, but to call the unit a work group (as opposed to a social group) is mistaken.

Alternatively, if the individual responses of members of the unit show strong task interdependence, shared responsibility, a sense of the unit as an intact social entity, and collective boundary management (all with low within-group variance), then finding high variance among group member responses to other group-related variables does not disqualify the unit as a work group. In fact, we may very well expect high within-group variation in regard to certain variables. For example, the group attribute of skill variety might vary with the organizational tenure of the group members (i.e., members with greater tenure may have greater skill variety than younger members as a result of more cross-training), but just because a unit has variation in skill variety doesn't make it less of a work group. Similarly, mixed responses to questions concerning the level of affect-based conflict within a group simply may reflect that members have varying ideas of when a disagreement becomes a conflict. Thus, having variance among other group-related variables is not a cause for concern in terms of defining a unit as a work

group. Of course, it does pose a problem for further group-level analysis of those variables (such as in examining relationships to group effectiveness), as a simple group average will be inappropriate in those cases.

An additional point should be made. Our discussion is focused on four elements that define a work group. It is not concerned with defining an effective work group. Whereas a researcher might posit, for example, that work groups with high levels of internal communication are more likely to be effective work groups, we are not suggesting that having low levels of internal communication (or any other group-related variable beyond the four elements we specify) detracts from a unit's "groupness." In other words, we make a distinction between variables that define a work group and variables that a researcher might propose characterize an effective work group. Our goal is to help ensure that studies of what enables work groups to be effective have as their unit of study true work groups.

Method

In this section, we use a previously collected data set to illustrate our arguments of how to determine if an organizational unit constitutes a work group.

Participants

Our sample is drawn from a research division of a large petrochemical corporation. The division performs centralized research for technical companies; its employees include many geologists, physicists, and engineers. Initial interviews at the firm indicated that several types of work units exist there, including functional departments, project teams, and management teams. While most of the units are involved in technical work (e.g., seismic information, thermal recovery, reservoir engineering), several constitute support units (e.g., accounting, purchasing, and information technology management).

This sample was chosen for this paper because we suspected that the project teams within it would rate highly in terms of groupness. Large functional departments, conceivably having multiple goals and obligations, in addition to partitioned tasks and sub-units, were expected to rate lower on the groupness continuum than the project teams, perhaps lying in the middle ranges. Management teams were considered as likely to qualify for work group status or not, depending on how they were structured and run; we expected to see some management teams that rated low on measures of groupness. Thus, the sample would allow us to determine if the four definitional elements are helpful in separating work groups from other collections of individuals at work.

Work Group Variables

We selected from the collected data set variables that were related to each of the four elements of the definition of a work group: task interdependence, shared responsibility, sense of being an intact social system, and collective boundary management. For both task interdependence and outreach responsibility, there was strong alignment between a variable in the data set and the definitional element. For the other two elements, there was no variable in the existing data set that exactly matched, so we selected proxy variables to demonstrate our approach. In the case of shared responsibility, the proxy variable was mission awareness. Questions related to mission awareness asked if the team knew its role, goals, and expected achievements. Reported unit size was used to operationalize the sense of being an intact social entity, on the grounds that agreement on unit size would indicate that the unit could distinguish itself as a separate organizational group. For comparison purposes, two additional group-related variables were examined: conflict and participation. Table 1 lists all the variables we employed and their operational measures. With the exception of unit size and outreach responsibility,

questions were answered using seven-point Likert-type scales. Outreach responsibility was a binary variable (zero if the unit did not have responsibility for the activity and one if it did). Unit size was written in by each respondent.

Procedure

Surveys were distributed to all 864 employees in the division, of whom 430 responded, yielding an individual response rate of 49.8%. Respondents represented 166 out of a possible 247 work units, for a unit response rate of 67.2%. Survey respondents were asked to identify their primary group by name and code from a list of all possible work units. They were further asked to identify the type of unit: department, project team, or management team. For the purposes of this paper, work units with responses from fewer than three members were eliminated from the data set, resulting in data from 269 individuals representing 60 work units: 40 departments, 13 project teams, and 7 management teams. The reported size of these units varied by type: departments ranged from 3 to 36 people, project teams from 3 to 11 people, and management teams from 3 to 14 people. The number of members reporting for a single unit was as high as nine for departments, five for project teams, and seven for management teams. About 30% of the respondents have a bachelor's degree; an additional 38% have completed some graduate work as well. Men comprise approximately 68% of the workforce in our sample.

Analysis

The responses for the members of each work unit were analyzed for within-group agreement using James et al.'s (1993) interrater agreement index, r_{wg} (renamed from the interrater reliability coefficient, which was introduced in their 1984 paper). In so doing, we built what Chan (1998) terms direct consensus models for each of the variables. Such models exist when the meaning of a higher level organizational construct is derived from achieving consensus

among the responses of lower-level units (in our case, the individual members of the work unit). Thus, we sought agreement among individual member perceptions about group-level constructs.

We considered employing other methods besides James et al.'s index, such as within-and-between-analysis (WABA; for examples of its use, see Dansereau, Alutto, & Yammarino, 1984, Van Aken & Kleiner, 1997, and Yammarino & Markham, 1992). We chose r_{wg} (despite the arguments that surround it, see Kozlowski & Hattrup, 1992, Yammarino & Markham, 1992, and Schmidt & Hunter, 1996 and 1989, in addition to the James et al. articles) because it allows us to examine responses and their variation on a group-by-group basis for a single variable at a time. Our intent is not to enter into the debate about r_{wg} or its use; rather, we simply desire to demonstrate our conceptual argument in this paper with the use of a statistical measure to examine the appropriateness of aggregation to the group level. Other measures for this purpose may be proposed and adopted in future research. We support the use of any statistical measure that permits group-level analysis, and do not mean by our use of James et al.'s measure to imply a preference for one method over another.

Results

Descriptive Statistics

Table 2 provides descriptive statistics, Pearson correlations, and Cronbach's alpha reliability coefficients for the work group variables. The high alpha coefficients indicate that the individual survey items meant to capture each variable were well-related, i.e. they captured the same construct.

Testing for Groupness

In testing for groupness, we needed to examine both within-group agreement as well as overall high values for each definitional element.

Within-group agreement.

George (1990) notes that $r_{wg} \geq .70$ indicates a reasonable degree of agreement among team members. Table 3 shows the number of work units having $r_{wg} \geq .70$. For example, 25 out of 40 departmental groups have $r_{wg} \geq .70$ for task interdependence. It is clear from the table that project teams best match the definition of work groups. Departments fared the next best, while management teams trailed in the percentage of units with strong internal agreement. Only for outreach responsibility (the variable aligned with collective boundary management) were the numbers for project teams similar to those for the other work units, with all values in the range of 50-60%. For this variable, management teams had the highest percentage of teams with low within-group variation.

The overall results are included with those for the individual variables because we contend that work groups ideally should match all four elements of the definition of a work group. Here again, project teams reflect higher satisfaction of the definitional elements, with 76.9% of the project teams having high member agreement on 3 out of 4 of the variables. By comparison, just over 40% of departmental groups and management teams display high agreement on 3 out of 4 of the elements. It is well worth noting that the percentages for all three types of teams drop steeply when the criterion is raised to having high agreement on all four variables. For example, less than a quarter of project teams have $r_{wg} \geq .70$ for all four variables.

It is interesting to note that less than half (46%) of the project teams had low within-group variation with regard to conflict and participation, two additional group variables we examined. These results lend support to our earlier argument; namely, a work unit's members can have disparate responses to variables commonly associated with group effectiveness without being considered less of a work group.

High values. In addition to verifying low within-group variation for the key elements, we must also check that the reported values for each associated variable are sufficiently high. If the variation is low and the values high, then the units can be considered work groups. Thus, the levels of task interdependence, mission awareness, and outreach responsibility are considered next. Reported group size is not considered, as we do not specify that a work unit must be within any given size range in order to qualify as a work group. For other data sets where a different variable (other than group size) is used to measure the intact social entity element, all four elements should be examined.

Figures 1-3 display histograms of the group-level values (taken as averages of the individual member responses) for those work units with $r_{wg} \geq .70$. Figure 1 displays the histogram for task interdependence, Figure 2 for mission awareness (the proxy variable for shared responsibility), and Figure 3 for outreach responsibility (the variable aligned with collective boundary management). Each of the figures indicates that these units lie predominantly on the high end of the scales, reflecting the units' high self-ratings for each of the variables. The few groups that tail off to the middle or lower end of the scales possibly lack sufficiently high values to be considered work groups. The majority of these units, however, pass both our requirements for being a work group: they have high internal agreement among members on the four key definitional elements, and they possess high levels of all elements except size, which we do not consider.

Discussion

We argue in this paper that researchers professing to study work groups should justify their claims based not on low within-group variation on the particular variables of interest in the study, but rather on low variation for variables that serve to define and identify a work group.

We further claim that low within-group variation must be coupled with sufficiently high values on these variables. Using Hackman's (1987) definition as our guide, we argue that work units demonstrating low within-group variation on and sufficient levels of task interdependence, shared responsibility, sense of being an intact social entity, and collective boundary management can appropriately be viewed as work groups.

Our analysis provides evidence of how these four definitional elements can be operationalized to verify that the work units we study qualify as work groups. In this regard, it proved rather successful in separating project teams from departments and management teams. A large percentage of project teams in our sample had low within-group variation and high values in three of the four elements that define a work group. For them, aggregation of individual responses to the group level for further analysis is appropriate. A much lower percentage of departments and management teams had such good results.

These results suggest that it is feasible, and we claim, desirable, for researchers to provide substantiation that the units they study are work groups. For groups where there is agreement and sufficient levels of the variables associated with the definition of a work group, we propose that it is appropriate to conceptualize these units as work groups. Researchers studying such groups can aggregate individual member responses for these variables to a group level for further analysis. Furthermore, prior work group research and theory (as distinguished from research on other types of groups) may be applicable to them. We argue that for units that have high within-group variation and/or low values for variables associated with the four definitional elements, it is questionable both to consider them work groups and to aggregate individual responses.

Quite possibly, a strict application of our procedure for all four elements of Hackman's definition may be too stringent a test in determining whether or not an organizational unit is a

work group. Recall that among project teams in our sample, over three-fourths of the teams had high within-group agreement on three out of four of the variables associated with the definitional elements. But less than a quarter of the teams had high agreement on all four of the variables. In such cases, researchers may decide, perhaps in consideration of qualitative data, that the units really ought to be considered work groups. They might find it prudent, however, to control for the four elements in further analysis, such as the building of predictive models of effectiveness.

Our results confirm a suspicion we have about many of the studies of management “teams.” As noted in Cohen and Bailey (1997), in studies of management units, the sample is often constructed by extracting the titles of managers from indices like Dunn & Bradstreet, then assuming that these top-echelon managers work together as a “team.” That management teams in our sample scored lower than functional departments in our tests for work groupness suggests that future researchers should provide much more thorough substantiation for their claims that the units they study are work groups. In the absence of such substantiation, we doubt that prior research and theory on work groups will be applicable to management units, nor do we believe that results of the management studies will have much relevance to work group research and theory.

Our approach begs the question of what to do in those instances where the units are indeed work groups, but display disparate within-group responses for other variables that are the major focus of the research. Recall that the project teams in our study had higher within-group variation for conflict and participation variables than they did for the four variables associated with the definition of a work group. The results for conflict and participation do not disqualify the project teams as work groups because these variables are not key to the definition of what constitutes a work group. They do suggest, however, that aggregation of individual responses to

the group level via simple averaging for these two variables is not appropriate, as the average may very well be unpopulated, and thus largely meaningless. In such cases, can a group level of analysis still be employed? We leave it to other researchers to further develop statistical measures for handling such situations, as that is not our focus or intent. We do caution, however, that we think such cases will arise frequently.

We further claim that just as work units cannot be disqualified from being work groups on the basis of high variation or low values on variables not related to the definition of a work group, they also cannot be labeled work groups on the basis of low variation and high values of other group-related variables. Such identification can only be made upon satisfaction of definitional requirements like the ones we propose in this paper.

We avoid specifying what constitutes a sufficiently high level of any of the key definitional elements. For example, we do not indicate how much task interdependence is required if a group of workers is to be considered a work group. Our primary reason for not doing so is quite simply because we cannot. Researchers must be familiar with the setting and content of the work to determine what level of interdependence is sufficient. Of course, we should all be skeptical if responses within a sample are concentrated at the low end of a scale. Similarly, when scores are at the high end, judgment is made easier. The task obviously becomes difficult only in the mid-ranges, and it is there that researchers must make their arguments based on qualitative assessments of particular situations.

There are several limitations to our analysis in this paper. First, our sample did not include all types of teams, as defined in Cohen and Bailey (1997). It did not include traditional work teams or parallel teams. We would expect such teams to perform as well as project teams on the definitional measures, but cannot provide evidence of that here. Second, our sample size is

somewhat small for project teams and management teams. The results would obviously be stronger if the sample were larger. Lastly, because we examined an existing data set, we were not able to define perfectly aligned variables to correspond to each of the four definitional elements. For example, we used reported team size as a proxy measure for a sense of being an intact social entity. Despite these limitations, our sample does point to greater “groupness” among some types of work units (i.e., project teams) than others (i.e., departments and management teams), and thus provides support for our argument.

In summary, we recommend that group researchers show evidence to establish that the units they study are work groups. For survey researchers, tests of the type we describe here are desirable. We suggest as an area of continued research the development of appropriate statistical procedures for arriving at a single group value for other group-related variables when individual responses are disparate in a unit that has low variance in the definitional ones. Additionally, it may be helpful to develop methods or guidelines for determining sufficient levels for the four elements that together serve to define a work group (e.g., how much task interdependence is sufficient to label a unit a work group?). We also recommend the development and consistent use of standard operational measures for these four elements to facilitate comparisons across work group studies.

References

Alderfer, C.P. (1977). Group and intergroup relations. In J.R. Hackman & J.L. Suttle (Eds.), Improving the quality of work life (pp. 227-296). Palisades, CA: Goodyear.

Cammann, C., Fichman, M., Jenkins, G. D., Jr., & Klesh, J.R. (1983). Assessing the attitudes and perceptions of organizational members. In S.E. Seashore, E.E. Lawler III, P.H. Mirvis, and C. Cammann (Eds.), Assessing Organizational Change: A Guide to Methods, Measures and Practices, pp. 71-138. New York: John Wiley and Sons.

Campion, M.A., Medsker, G.J., & Higgs, A.C. (1993). Relations between work group characteristics and effectiveness: Implications for designing effective work groups. Personnel Psychology, 46, 823-850.

Cartwright, D. & Zander, A. (1968). Groups and group membership: Introduction. In (Eds.), Cartwright and Zander, Group Dynamics. New York: Harper and Row.

Chan, D. (1998). Functional relations among constructs in the same content domain at different levels of analysis: A typology of composition models. Journal of Applied Psychology, 83(2), 234-246.

Cohen, S.G., & Bailey, D.E. (1997). What makes teams work: Group effectiveness research from the shop floor to the executive suite. Journal of Management, 23(3), 239-290.

Cohen, S.G. & Ledford, G.E. (1994). The effectiveness of self-managed teams: A quasi-experiment. Human Relations, 47(1): 13-43.

Cohen, S.G., Ledford, G.E., & Spreitzer, G.M. (1996). A predictive model of self-managing work team effectiveness. Human Relations, 49(5): 643-676.

Dansereau, F., Alutto, J.A., & Yammarino, F.J. (1984). Theory testing in organizational behavior: The variant approach. Englewood Cliffs, NJ: Prentice-Hall.

George, J.M. (1990). Personality, affect, and behavior in groups. Journal of Applied Psychology, 75(2), 407-446.

George, J.M. (1995). Leader positive mood and group performance: The case of customer service. Journal of Applied Social Psychology, 25(9): 778-794.

George, J.M. & Bettenhausen, K. (1990). Understanding prosocial behavior, sales performance, and turnover: A group-level analysis in a service context. Journal of Applied Psychology, 75(6): 698-709.

Gibson, C.B., Randel, A.E., & Earley, P.C. (in press). Understanding group-efficacy: An empirical test of multiple assessment methods. Group and Organization Management.

Guzzo, R.A., & Dickson, M.W. (1996). Teams in organization: Recent research on performance and effectiveness. Annual Review of Psychology, 47, 307-338.

Hackman, J.R. (1987). The design of work teams. In J.W. Lorsch (Ed.) Handbook of Organizational Behavior, pp. 315-342. Englewood Cliffs, NJ: Prentice-Hall.

Hallam, G. & Campbell, D. (1994). Test manual for the Campbell-Hallam team development survey. NCS Assessments: Minnetonka, MN.

James, L.R., Demaree, R.G., & Wolf, G. (1993) r_{wg} : An assessment of within-group interrater agreement. Journal of Applied Psychology, 78(2), 306-309.

James, L.R., Demaree, R.G., & Wolf, G. (1984). Estimating within group interrater reliability with and without response bias. Journal of Applied Psychology, 69, 85-98.

Kirkman, B.L., Tesluk, P.E., & Rosen, B. (1998, August). Comparing the aggregation of individual responses versus team consensus ratings in measuring team level data: If you want team level data, shouldn't you ask the team? Paper presented at the meeting of the Academy of Management, San Diego, CA.

Klein, K.J., Dansereau, F. & Hall, R.J. (1994). Levels issues in theory development, data collection, and analysis. Academy of Management Review, 19(2), 195-229.

Lewin, K. (1948). Resolving social conflicts. New York: Harper.

Murrell, A.J., & Dietz, B. (1992). Fan support of sport teams: The effect of a common group identity. Journal of Sport and Exercise Psychology, 14, 28-39.

Pearce, J.A. II & Ravlin, E.C. (1987). The design and activation of self-regulating work groups. Human Relations, 40(11): 751-782.

Pearson, C.A.L. (1992). Autonomous workgroups: An evaluation at an industrial site. Human Relations, 45(9): 905-936.

Rousseau, D. (1985). Issues of level in organizational research: multi-level and cross-level perspectives. In L.L. Cummings and B.M. Staw (Eds.), Research in Organizational Behavior, 7, pp. 1-37. Greenwich, CT: JAI Press.

Schmidt, F.L., and Hunter, J.E. (1996). A critical examination of r_{wg} as an index of interrater agreement. Unpublished working paper. College of Business, University of Iowa City, Iowa.

Schmidt, F.L., and Hunter, J.E. (1989). Interrater reliability coefficients cannot be computed when only one stimulus is rated. Journal of Applied Psychology, 74(2), 368-370.

Sundstrom, E., DeMeuse, K.P. & Futrell, D. (1990). Work teams: Applications and effectiveness. American Psychologist, 45, 120-133.

Van Aken, E.M., & Kleiner, B.M. (1997). Determinants of team effectiveness for cross-functional organizational design teams. Quality Management Journal, 4(2), 51-79.

Yammarino F.J. & Dubinsky, A.J. (1990). Salesperson performance and managerially controllable factors: An investigation of individual and work group effects. Journal of Management, 16(1): 87-106.

Yammarino, F.J., & Markham, S.E. (1992). On the application of within and between analysis: Are absence and affect really group-based phenomena? Journal of Applied Psychology, 77(2), 168-176.

Table 1

Operational Measures from Existing Data Set

Element of Work Group Definition	Variable from Existing Data Set	Operational Measure
Task interdependence	Task interdependence	I cannot accomplish my tasks without information or materials from other members of my team. Other members of my team depend on me for information or materials needed to perform their tasks. Within my team, jobs performed by team members are dependent on one another. (each item on a 7-point Likert-type scale)
Shared responsibility	Mission awareness	My team knows what we are supposed to accomplish. My team has clear goals it is trying to achieve. My team is clear about what its role is. (each item on a 7-point Likert-type scale)
Intact social entity	Size of the unit	How many members are in your team? (blank line for write-in response)
Boundary management	Outreach responsibility	Do members have significant responsibility for maintaining contact with: (a) individual contributors outside the team, (b) other teams, (c) key suppliers to the team, (d) the team's customers/clients, (e) upper level managers, or (f) colleagues outside the company? (each item checked or not, coded as 1 or 0)
Additional Group-Related Variables		Operational Measure
	Conflict	My team attempts to resolve disagreements in a constructive manner. My team proposes a reasonable approach to resolve a disagreement. My team encourages frank and open discussion of a disagreement. (each item on a 7-point Likert-type scale)
	Participation	We encourage each other to participate in making important decisions. We encourage people to speak up when they disagree with a decision. (each item on a 7-point Likert-type scale)

Note. Operational measures for task independence were drawn from Campion, Medsker, and Higgs (1993). Measures for participation were taken from Cammann, Fichman, Jenkins, and Klesh (1983). All other measures were created for the original data set.

Table 2.

Descriptive Statistics, Variable Correlations, and Alpha Reliability Coefficients

Variable	Mean	s.d.	1	2	3	4	5	6
1. Size	8.51	4.61	-					
2. Conflict	5.13	1.25	-.12	.85				
3. Participation	5.27	1.30	-.04	.81	.79			
4. Task Interdependence	4.93	1.18	.04	.22	.17	.71		
5. Mission Awareness	5.17	1.36	-.19	.53	.38	.17	.87	
6. Outreach Responsibility	.67	.29	.06	.04	.02	-.02	-.10	.77

Note. Cronbach's alpha coefficients appear in bold on the diagonal.

Table 3.

Units with $r_{wg} \geq .70$

Variable	Departmental groups (n=40)	Project teams (n=13)	Management teams (n=7)
Group-definition variables (element)			
1. Task interdependence (task interdependence)	25 (62.5%)	12 (92.3%)	4 (57.1%)
2. Mission awareness (shared responsibility)	19 (47.5%)	10 (76.9%)	2 (28.6%)
3. Reported team size (intact social entity)	26 (65%)	10 (76.9%)	3 (42.9%)
4. Outreach responsibility (boundary management)	21 (52.5%)	7 (53.8%)	4 (57.1%)
Overall (3 out of 4 elements)	17 (42.5%)	10 (76.9%)	3 (42.9%)
(4 out of 4 elements)	7 (17.5%)	3 (23.1%)	1 (14.3%)
Additional group-related variables			
1. Conflict	21 (52.5%)	6 (46.2%)	5 (71.4%)
2. Participation	24 (60%)	6 (46.2%)	4 (57.1%)

Note. Number in parentheses is percentage of total number of groups within that category having $r_{wg} \geq .70$.

Figure 1. Task interdependence for work units with $r_{wg} \geq .70$

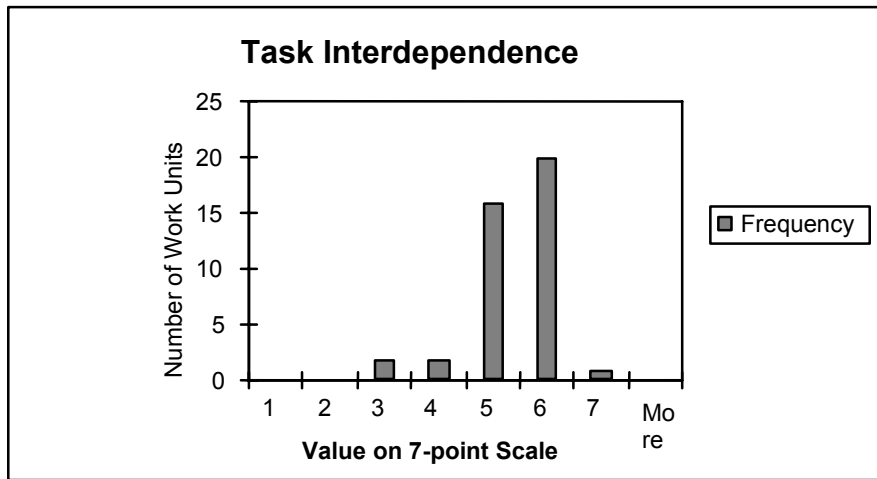


Figure 2. Mission awareness for work units with $r_{wg} \geq .70$.

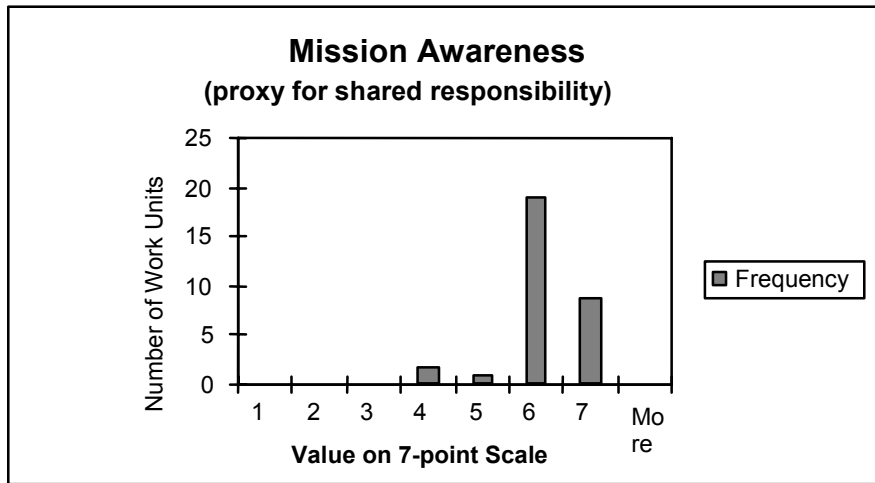


Figure 3. Outreach responsibility for work units with $r_{wg} \geq .70$.

