AFRL-HE-AZ-TR-2000-0128

UNITED STATES AIR FORCE RESEARCH LABORATORY

DEVELOPMENT AND TEST OF A THEORY OF WORK TEAM PRODUCTIVITY

John E. Sawyer

Department of Business Administration
University of Delaware
Newark, DE 19716

Robert D. Pritchard

Texas A&M University
College Station, TX 77843-4235

Winston R. Bennett, Jr.

Air Force Research Laboratory
Warfighter Training Research Division
6030 South Kent Street
Mesa, AZ 85212-6061

William R. Latham

University of Delaware
Newark, DE 19716

July 2001

AIR FORCE RESEARCH LABORATORY
HUMAN EFFECTIVENESS DIRECTORATE
WARFIGHTER TRAINING RESEARCH DIVISION
6030 South Kent Street
Mesa AZ 85212-6061

NOTICES

The views expressed in this report are those of the authors and do not necessarily reflect official views of the US Air Force or the Department of Defense.

Using Government drawings, specifications, or other data included in this document for any purpose other than Government-related procurement does not in any way obligate the US Government. The fact that the Government formulated or supplied the drawings, specifications, or other data, does not license the holder or any other person or corporation, or convey any rights or permission to manufacture, use, or sell any patented invention that may relate to them.

The Office of Public Affairs has reviewed this report, and it is releasable to the National Technical Information Service, where it will be available to the general public, including foreign national.

This report has been reviewed and is approved for publication.

//Signed//
WINSTON R. BENNETT, JR.
Project Scientist

//Signed//
DEE H. ANDREWS
Technical Director

//Signed//
JERALD L. STRAW, Colonel, USAF
Chief, Warfighter Training Research Division

Requests for copies of this report may be directed to:

Defense Technical Information Center
8725 John J. Kingman Road, Suite 0944
Ft. Belvoir, Virginia 22060-6218
http://stinet.dtic.mil

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. REPORT DATE *(DD-MM-YYYY)*	**2. REPORT TYPE**	**3. DATES COVERED** *(From - To)*
July 2001	Final	Jun 98 – Jun 00

4. TITLE
Development and Test of a Theory of Work Team Productivity

5a. CONTRACT NUMBER

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER
62202F

6. AUTHOR(S)
John E. Sawyer
Robert D. Pritchard
Winston R. Bennett, Jr.
William R. Latham

5d. PROJECT NUMBER
1123

5e. TASK NUMBER
A2

5f. WORK UNIT NUMBER
19

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

Department of Business Administration
University of Delaware
Newark, DE 19716

Texas A&M University
College Station, TX 77843-4235

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)
Air Force Research Laboratory
Human Effectiveness Directorate
Warfighter Training Research Division
6030 South Kent Street
Mesa AZ 85212-6061

10. SPONSOR/MONITOR'S ACRONYM(S)
AFRL

11. SPONSOR/MONITOR'S REPORT NUMBER(S)
AFRL-HE-AZ-TR-2000-0128

12. DISTRIBUTION / AVAILABILITY STATEMENT

Approved for public release; distribution is unlimited.

13. SUPPLEMENTARY NOTES
Air Force Research Laboratory Technical Monitor: Dr. Winston R. Bennett, Jr., AFRL/HEA, (480)988-6561, X-297; DSN 474-6297

14. ABSTRACT
As noted in A Theory of Behavior in Organizations , by J.C. Naylor, R.D. Pritchard, and D.R. Ilgen's (1980), we extend their theory of behavior in organizations to explain work team productivity, and test it using time-series analysis on data from a large-scale study of teams. Priority scores for 37 indicators of productivity across five work teams over 23 months were used to predict month-by-month changes in productivity for each of the 37 team products. The results show that team productivity improvements can be explained by feedback including priority scores derived from nonlinear contingency functions of the product indicators. Furthermore, teams which initially performed more poorly benefited the most from the priority score feedback. Goal setting positively affected productivity gain and did so consistently across work teams, after the effects of priority feedback and the interaction of work team with priority feedback were controlled. Implications for team performance strategies are discussed.

15. SUBJECT TERMS
NPI Theory; Team Behavior; Team Performance; Teams; Time Series Analysis; Work Team Productivity

16. SECURITY CLASSIFICATION OF:			**17. LIMITATION OF ABSTRACT**	**18. NUMBER OF PAGES**	**19a. NAME OF RESPONSIBLE PERSON**
					Dr. Winston Bennett, Jr.
a. REPORT	**b. ABSTRACT**	**c. THIS PAGE**			**19b. TELEPHONE NUMBER** *(include area code)*
UNCLASSIFIED	UNCLASSIFIED	UNCLASSIFIED	UL	26	(480) 988-6561, X-297

Standard Form 298 (Rev. 8-98)
Prescribed by ANSI Std. Z39.18

PREFACE

The authors thank Thomas Becker and Jerry Griffith for their careful review and comments on a previous version of this manuscript.

Correspondence regarding this article should be addressed to John E. Sawyer, Department of Business Administration, University of Delaware, Newark, Delaware 19716. Electronic mail may be sent via internet to: sawyerj@college.be.udel.edu.

DEVELOPMENT AND TEST OF A THEORY OF WORK TEAM PRODUCTIVITY

Productivity in most modern work teams results from a complex inter-relationship of multiple demands and multiple measures of performance. Work teams seldom have unlimited resources. In today's downsized organizations, work teams are typically operating with minimal staff. To be competitive, work teams must adhere to strict time and cost constraints. Thus, to be effective, work teams must target their resources (skills, abilities, effort, time, equipment, financial, etc.) at those activities that will result in the greatest increase in performance. These complexities and resource constraints require information that can help teams best apply their resources to important task accomplishments.

This paper applies the theory of behavior in organizations proposed by Naylor, Pritchard, and Ilgen (1980)(from here on referred to as NPI Theory) to develop a theoretical basis for understanding team performance. We will first summarize the relevant aspects of NPI theory, then relate the theory to team performance. We follow the theory development by summarizing the methods applied by Pritchard, Jones, Roth, Stuebing, & Ekeberg(1988; 1989)and detailing an alternative time series analysis which allows us to test the application of NPI theory to team performance. Our analysis contributes beyond that previously published by applying appropriate time series methods to resolve previously unanswered questions about the role of priority feedback in team performance.

Elements of "A Theory of Behavior in Organizations"

Naylor et al. (1980) suggested that individuals behave according to a strategy having to do with a series of acts. That strategy, they suggested, is formulated on the basis of patterns of multiple contingencies relating strings of acts to products, evaluations, and ultimately, to outcomes. In NPI Theory a contingency is "a tendency for two attributes or characteristics to exhibit some type of formal relationship" (p. 286). For example, a product-to-evaluation contingency is "the relationship existing between the amount of a particular product perceived by an evaluator . . . and the evaluation of the product by the evaluator" (p.286). In NPI theory, the product-to-evaluation contingencies[1] of others reflects ones "perceptions of (a) what products are evaluated and (b) the relative importance of the products in the evaluation system of the evaluator" (p. 40).

In NPI theory, the utility function for a product is a composite of the contingencies relating that product to outcomes (i.e., product-to-evaluation and evaluation-to-outcome contingencies) and the valence of the outcomes. Thus, the utility function for each product reflects the individual's perception of the degree to which variations in the product result in variations in anticipated valued outcomes. The motivational force for producing a given product is dependent on the slope of the utility function for that product.

The implication of defining motivational force as the slope of the utility function is that a person will commit additional resources to a given product for which <u>change</u> in the amount of resources committed (ΔC)will result in the greatest change in utility (ΔU) of that product. It is important to note that the theory focuses attention on change in resource commitment (in addition to resources already committed). The importance of both the current amount of commitment and size of the change in commitment is that with any nonlinear function, the "ratio $\Delta U/\Delta C$ will be different depending upon (a) where one starts; and (b) the size of the interval over which the slope is computed" (Naylor et al., 1980). Thus the theory is dynamic. Many actions in work settings become routine or automatic. The theory may be used to explain how one got to a given point, but more importantly allows one to plan strategies to reallocate resources to products that will provide the greatest improvement in overall effectiveness.

Naylor et al. (1980) cite substantial research regarding the ability of people to use nonlinear functions in decision making. Sawyer (1991) found that individuals effectively learned nonlinear contingencies. Furthermore, when individuals were asked to allocate personal time to tasks represented by different functional relations, they were able to use their knowledge of nonlinear functional relations to optimally allocate time (Sawyer, 1990). Sawyer's (1990) findings suggest that mechanisms that help individuals learn what products are valued by evaluators and product-to-evaluation contingencies for those evaluators will lead to more optimal allocation of resources such as time and personal effort. In the following section we expand the NPI theory to include work teams and elaborate conditions that may enhance team performance.

Application of NPI Theory to Team Strategy

The resources teams have to commit to tasks include time, effort, skills, and abilities of their members, as well as financial and material resources. In complex tasks, team members may contribute different skills and perform different functions. In addition to tasks related to daily work output, there are tasks related to development of member and team skills. Activities that keep the team abreast of new technical developments and work processes are valued contributions to the team's overall effectiveness, along with the daily work outputs of the team. The strategic tasks of the team then include both the use of resources to complete daily tasks in an efficient and effective manner, and appropriate allocation of team resources (time, etc.) to development of the work capabilities of the team. Thus, daily work outputs as well as team skill maintenance can be viewed as products in the framework of NPI theory.

A critical difference between individuals and teams stems from the fact that team learning is different from individual learning. Much of individual behavior is based on tacit knowledge (Sternberg & Wagner, 1993; Sternberg, Wagner & Okagaki, 1993) or specific job knowledge (Schmidt & Hunter, 1993) held by the individual. As individuals receive feedback, they modify their perceptions of contingencies to more accurately reflect the form of the contingency (Sawyer, 1991). These changes in perception are then reflected in future actions (Sawyer, 1990). However, debate in the realm of organizational learning (Schneider & Angelmar, 1993) and team mental models (Klimoski & Mohammed, 1994) suggests that we cannot infer team or organizational processes from individual processes.

In translating learning of individuals to that of teams, we may consider the arguments of Argyris and Schön (1996) regarding organizational learning. Members of teams, as do members of organizations, construct their own cognitive representations of task requirements, which include inferences about causal relations between actions and outcomes (i.e., NPI theory contingencies). Individual representations by skilled and experienced members may hold complete representations of their individual tasks, however, their representations may lack specifics of tasks for which they are not personally responsible, or tasks for which they perform only a portion of the task. Considering that work team productivity includes a complex set of interrelations, individual representations may be least likely to reflect task characteristics that are not in the control of single individuals, but are products of those complex interrelations. Argyris and Schön (1996) suggested that even in face-to-face contact, as in work teams, private images of the organization often diverge. Thus, for the team to develop a comprehensive set of contingencies that represent effective team functioning, external references must be developed to guide private adjustments to obtain common (or at least compatible) representations of team tasks.

While many mechanisms exist to aid in forming team understanding of coordinated tasks (Klimoski & Mohammed, 1994), one mechanism has been developed which is consistent with the NPI theory notion of contingencies and allows us to apply a judgment-based model such as NPI theory to teams. Pritchard, Jones, Roth, Stuebing, and Ekeberg (1988) developed the Productivity Measurement and Enhancement System (ProMES) which can have the effect of guiding team members to explicitly recognize and share the contingency forms among the team members. The system seeks to define productivity in terms of organizational objectives. The system is bottom-up, in that a design team composed of representative members of the work unit, supervisors, and facilitators develop the measurement system. Once the team agrees on the specifics of the system, they take it to top management for discussion and approval. Once management approves the system, it can be considered to reflect a common understanding of what constitutes productivity for the work group.

The system of indicators and contingencies reflects a map of the essential relationships between products and evaluation of the work unit's productivity. These are the important elements of NPI theory product-to-evaluation contingencies. Specifically, products are identified, measures of those products are specified, and how those products will be evaluated is defined. Furthermore, these are specified at the group level of analysis because they are defined by and for the work unit, as opposed to a mere composite of product-to-evaluation contingencies specified for individuals within the work unit. Additionally, the system relates performance on all indicators to a single evaluation of overall effectiveness, a desirable trait of productivity measurement systems (Pritchard, 1990) and a key tenant in the evaluation system of NPI Theory (cf. chapter 2, Naylor et al., 1980).

Relating back to NPI theory, work units should guide their commitment of resources to those products for which the greatest gain in effectiveness can be achieved with a given amount of effort. In the ProMES system, the indicators are scaled so that they each contain approximately the same number of increments designed to reflect similar amounts of effort or resources for each indicator. In development of priorities from the ProMES system, work teams determine the amount of gain in effectiveness that could be obtained from one unit gain on each indicator. Thus, the team can easily determine which products to improve performance to obtain the greatest gain in overall effectiveness.

Pritchard et al. (1989) argued that the "system ... provided considerable information for developing productivity enhancing strategies" (p. 105). Work teams can best benefit from the productivity measurement system by using the information provided at the indicator level. The priority score for a given indicator is the change in effectiveness that can be obtained from improvement on that indicator within the next measurement period. These priority scores depend on the team's current level of performance and the form of the contingency function for each indicator. Because the relationships between the units of measurement and effectiveness scores for most indicators are nonlinear, priority scores vary depending upon the level of performance for each indicator[2]. Thus, priority scores allow the team to compare products in terms of potential gain relative to current performance.

Testing the NPI Theory in Work Teams

Information about the evaluated products in a team's performance system could be used in a variety of ways. One possibility is that a very simple strategy of applying effort toward indicators of performance on which the team had the most to gain is used. A comparison activity, such as that suggested by control theory (Klein, 1989), would lead teams to improve on those indicators that have the greatest overall potential gain. The simple strategy would then be to merely identify which indicators the team is performing furthest from the maximum and concentrate effort on those indicators.

If the above is sufficient to explain observed increases in performance, the ProMES system is unnecessarily complicated and the extension of NPI theory to work teams is not validated. However, since priority scores capture the essence of the nonlinear contingencies in NPI theory, if work teams are capable of using the priority scores to allocate resources toward products, priority scores should predict subsequent improvement on specific products. If work groups are using the priorities based on nonlinear contingencies as proposed by Naylor et al. (1980) and Pritchard (1990), we should find that such priorities account for significant variance in improvement in addition to the maximum potential gain on individual indicators of productivity. Thus, we test hypothesis 1:

> Hypothesis 1: Feedback of priority data by indicator will be positively associated with productivity improvement on specific indicators after controlling for the maximum possible amount of gain in each indicator.

It should be noted that controlling for maximum possible gain in each performance indicator, as proposed in hypothesis 1, allows us to rule out regression to the mean effects that might otherwise explain increasing performance on products for which performance is initially low.

Hypothesis 1 may hold for only some teams receiving priority feedback by indicator. We suggest that many mechanisms may make priority feedback more or less useful to teams. For instance, mutual adjustment, effective communication, and interpersonal coordination among highly functioning teams may make explicit systems such as ProMES unnecessary (Klimoski & Mohammed, 1994). Values, action strategies, and assumptions embedded in an organization’s “theory-in-use” may remain tacit rather than explicit (Argyris & Schön, 1996). Thus, some teams may not need the additional strategy information imbedded in the priority feedback. Poorly performing teams may be most helped by the strategy provided by priority feedback because they have not formed other effective systems. Thus, to adequately test the effects of priority feedback, we must control for team differences. Thus, we propose hypothesis 2:

> Hypothesis 2: There will be a greater association between priority feedback and productivity on individual products among work units that perform more poorly initially than initially high-performing work groups.

Previous research has suggested that formal goal setting is most effective when proper task strategies are in place (Wood & Locke, 1990). Recent studies of goal setting involving both individuals and teams have identified the critical role of work strategies. For example, recent theoretical developments in the relationship of goal setting to task strategies (Earley, Connoly, & Ekegren, 1989; Earley, Lee & Hanson, 1990; Wood & Locke, 1990) have shown that goal setting works by guiding the development of task strategies. Studies of the use of goal setting in teams (Mitchell & Silver, 1990; Weingart, 1992; Weingart & Weldon, 1991; Weldon, Jehn, & Pradhan, 1991; Wood & Locke, 1990) have shown that goal setting has strongest effects when appropriate strategies are developed. We suggest that the development of product contingencies and the use of priority feedback establish the necessary foundation for goal setting to be most effective. Therefore, we hypothesize that once maximum gain, the interaction of work team with indicator priority feedback, and time series effects are controlled:

> Hypothesis 3: Formal team goal setting will affect a substantial improvement in effectiveness on productivity indicators.

We do not expect an interaction of work teams with goal setting. This is based on the work of Wood, Mento, and Locke (1987), which demonstrated that differences in goal effects across studies can be accounted for by task strategies. By controlling for work group, priority feedback, and work group by feedback interactions, any variation in strategy should be controlled. Thus, we expect no further interaction with work group. We include interactions of work groups with goal setting in our initial analysis to confirm our assumption, but since we expect the null, we do not specify a hypothesis regarding this lack of effect.

The literature suggests a general consensus that incentives enhance goal commitment (Riedel, Nebeker, & Cooper, 1988; Wright, 1992). If the interaction of work teams with priority feedback, formal goal setting, and time series effects are controlled we hypothesize that:

> Hypothesis 4: The addition of group incentives will lead to significant improvement in effectiveness on productivity indicators.

Again, we do not expect an interaction of work team with incentives. We include interactions of work groups with incentives in our initial analysis to confirm our assumption, but since we expect the null, we do not specify a hypothesis regarding this lack of effect.

Application of Time Series Analysis

Pritchard et al. (1988) used the ProMES to provide feedback (including priority information) regarding the various indicators of productivity, to explicitly set goals in terms of overall team effectiveness, and to provide incentives for productivity improvement. In the analyses reported by Pritchard et al. (1988), the percentage of maximum possible gain in overall effectiveness (summing across all indicators for a shop) achieved during each period relative to the baseline period was 50% during the feedback period, 75% during the goal setting period, and 76% during the incentive period. Because the interventions were applied sequentially and additively (i.e., once a treatment was begun, it continued to the end of the experimental period), one would be tempted to infer that feedback alone had a greater impact on team effectiveness (50%) than did the addition of goal setting (+25%), and incentives had trivial effects (+1%). Pritchard et al. (1988) discussed at length possible interpretations of their findings and concluded that "goal setting added somewhat to feedback, but not a great deal, and that incentives did not add further" (p 353). This is contrary to Locke and Latham's (1990) contention that feedback cannot affect motivation apart from goal setting. A number of logical explanations were offered by Pritchard et al. (1988) including learning effects, ceiling effects, and possible informal goal setting by teams prior to the formal goal-setting period.

Another possible explanation is that time series effects may have exaggerated the apparent effects of feedback, and/or may have masked the effects of goal setting. Recent research on time series effects (e.g., Bergh, 1993a, 1993b) makes it prudent to reanalyze the data previously reported by Pritchard, et al. (1988). Bergh (1993a, 1993b) has demonstrated the need to explore time series effects in management research: without appropriate time series analysis, it is impossible to confidently offer interpretations of effects observed over time. Aggregating time series data into before and after periods and conducting Analysis of Variance (ANOVA) tests, as done by Pritchard et al. (1988), ignores the time characteristics in the data that may result in misleading or biased interpretations (Bergh, 1993a).

Additionally, Pritchard et al. (1988; 1989) did not conduct their analysis at the indicator level. Their analysis was conducted after effectiveness scores were summed across all indicators. Thus, the previously reported analyses do not allow any conclusions about the association between priority scores for a given indicator of performance and subsequent improvement on that indicator. Thus we cannot infer that work teams actually used these priorities to direct their efforts to improve indicators from which the greatest increase in effectiveness could be achieved. While Pritchard et al. (1988, 1989) argued that this is likely what happened, they admitted that their discussion of motivational factors were "speculative and only supported by anecdotal data" (Pritchard et al., 1989, p. 104).

In the present paper we reanalyze the Pritchard et al. (1988) data using appropriate time series methods. Our reanalysis of the Pritchard et al. (1988) data is motivated by two important concerns. First, we believe this data set provides an excellent opportunity to test the theory of team performance based on the Naylor et al. (1980) theory of behavior in organizations. The previous analysis does not allow inferences about the use of priority feedback because that analysis did not: (a) control for maximum possible gain; (b) control time series effects, causing uncertainty about the effects of feedback, goal setting and incentives; and (c) investigate the association between the actual priority scores for each indicator and subsequent improvement in the specific indicator. Our analysis directly addresses the cautions and recommendations of Bergh (193a; 1993b) concurrently with testing the applicability of NPI theory to team performance.

METHOD

Archival data from the ProMES research conducted with five maintenance and materials storage and distribution units of the United States Air Force (reported by Pritchard et al., 1988) were organized by indicator. Procedures for developing and implementing the process are described in detail in Pritchard et al. (1988; 1989)[3]. The sample and the procedures developing ProMES contingencies are summarized below followed by the procedures for the time series analysis.

Method for Development and Assessment of ProMES

The sample for the study included a Communications and Navigation (Comm/Nav) maintenance section and four sections of the Material Storage and Distribution Branch (including Receiving, Storage and Issue, Pickup and Delivery, and Inspection units) of the United States Air Force at a base in the southwest United States. The number of personnel in each unit ranged from 7 (Inspection) to 35 (Comm/Nav). Total number of personnel involved was 83. While there was some turnover within each unit, identity of the unit and the members involved in developing and maintaining the ProMES system within each unit remained stable over the 23-month period of the study. All of the data for this analysis were taken at the work unit level of analysis, so the small level of turnover in membership experienced within the units was not a concern.

Four steps were followed in constructing the ProMES in each of the target shops. These included: (a) identifying salient products, (b) developing indicators of these products, (c) establishing contingencies, and (d) putting the system together as a feedback system. The first three steps were performed by a group representing each shop, composed of supervisors and representative incumbents. The fourth step was conducted by the experimenter on the basis of contingencies developed by the shop and measurements taken of performance on each indicator.

Step 1: Identify Products

Products are the activities or objectives that the shop is expected to accomplish. Supervisors and representative incumbents met to identify the salient products that make up the essential activities and objectives of the shop.

Step 2: Develop Indicators

An indicator is a measure of how well the shop is generating the specific product. The group of supervisors and incumbents were asked to think of things they would use to show how well they were generating each of their products. Products included performance output as well as team maintenance activities (e.g., training and technical qualification tasks). There were one or more indicators for each product. The indicators identified for the five shops in this research are presented in Table 1.

Table 1. List of Indicators by Shop

Shop	Indicator
Comm/Nav	
	C1 %Bounces
	C2 QA Inspections
	C3 Number of units Awaiting Maintenance (AWM)
	C4 Units Awaiting Parts (AWP)
	C5 %Demand Met
	C6 Specialty Training Standard (STS) tasks complete
	C7 % Qualification Tasks Comp: C
	C8 % Qualification Tasks Comp: N
	C9 Scheduled Training Tasks Overdue [a]
	C10 Mobility Equipment Overdue [a]
	C11 Precision Measurement Equipment Lab. Units Overdue
	C12 349 (Maintenance Data Collection Record) Errors
	C13 Missed Appointments
MS&D: Storage and Issue	
	S1 Warehouse Refusals in Wrong Location
	S2 # Findings
	S3 % Cleared Off R36 list
	S4 Priority 2 Items to issue (Mins)
	S5 Priority 3 Items to issue (Mins)
	S6 Priority 4 Items to issue (Hrs)
	S7 # Repeat Findings [a]
MS&D: Receiving	
	R1 Inchecking Errors/1000
	R2 Priority 2 Items to Pickup (Mins)
	R3 Priority 4 Items to Pickup (Hrs)
	R4 Warehouse Refusals in Receiving
	R5 Average # Delinquent Rejects
MS&D: Pickup and Delivery	
	P1 Delinquent Pickups
	P2 ISU's to Customer (Mins)
	P3 Priority 2 DOR's to Customer (Mins)
	P4 Priority 3 Items to Customer (Mins)
	P5 Priority 4 Items to Customer (Hrs)
	P6 # Delivered Wrong Location [a]
	P7 Vehicle Inspection Score
	P8 # Reportable Accidents [a]
	P9 # Nonreportable Accidents
MS&D: Inspection	
	I1 Average # Local Incoming Items
	I2 Average # Dated Items From Depot
	I3 Average # Unidentified Items
	I4 Average # Aircraft Parts left
	I5 Average # Functional Checks left
	I6 Average # Suspect Items left
	I7 # Late Monthly Inspections [a]
	I8 # Returns From DPDO
	I9 % TCTOS Checked [a]
	I10 # Rods From Other Bases
	I11 Average # Off Base Shipments

[a]Indicator was removed from the analysis because it was at maximum effectiveness at the beginning and throughout the observation period.

Step 3: Establish Contingencies

A contingency was established for each indicator. The term contingency is used here to refer to the level of evaluation of an outcome that is contingent on the amount of that outcome. The level of evaluation is the effectiveness of the shop when a given amount of the outcome is produced.

To establish the contingency for each indicator, the group of shop supervisors and incumbents were asked first to identify the maximum (or best) possible amount of the outcome and the minimum (worst) feasible amount of the outcome. Second, the group was asked to determine the neutral point for the indicator; that is the point at which the amount of the outcome is neither good nor bad. That amount for the given indicator was located at zero on the effectiveness scale. In Figure 1, for the indicator "units awaiting maintenance," in the Comm/Nav shop, 60 units has been determined to be neutral and has been placed at 0 effectiveness points.

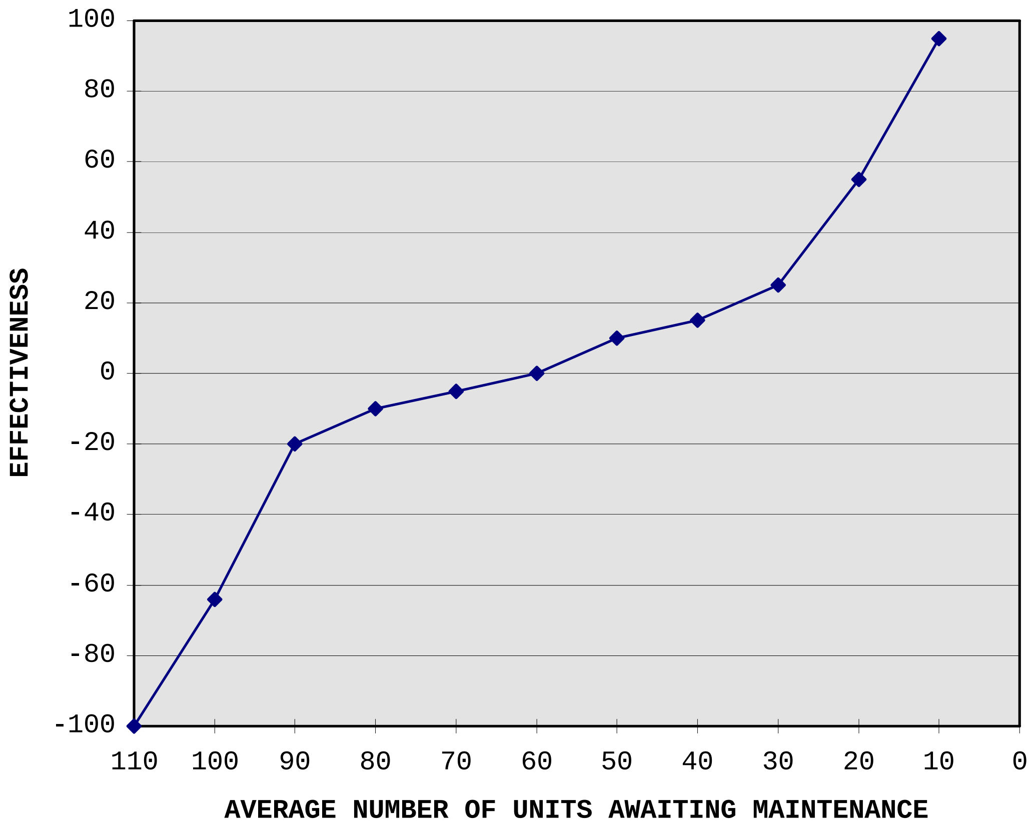

Figure 1. Sample Contingency for the Communication and Navigation Team

Then the group determined the effectiveness level of the minimum and maximum amount of output. First, all of the indicators for the shop were listed along with the maximum outcome level for each indicator. The group then rank ordered these maximums in terms of the contribution of each to the overall effectiveness of the unit. The group discussed their rankings until a consensus was reached. The maximum with the highest importance rank was given an effectiveness value of +100. The group was then told to rate the other maximums as percentages of the +100 maximum (e.g., if the maximum of a given indicator was only half as important to the effectiveness of the shop as the most important maximum, they rated it an effectiveness value of +50. An analogous process was done for the minimum values of each indicator, except the most important (worst) minimum was not constrained to be an effectiveness value of -100. It was given a negative value the group thought was appropriate (i.e., its importance to the effectiveness of the shop in proportion to the importance of the most important maximum).

Finally, the intermediate points were assigned effectiveness values by the group. Group discussion continued until consensus was reached. In Figure 1 the contingency for the indicator "units awaiting maintenance" established by the Communications and Navigation shop group is illustrated. Its maximum (best) level of 10 units was rated 95 effectiveness points. Its minimum (worst) level of 110 units was rated -100 effectiveness points. Each intermediate 10 unit was given an effectiveness rating to produce the contingency function illustrated in Figure 1. This process was repeated for each indicator within the shop.

Step 4: Put the System Together

First the indicator data were collected for a given month. Based on the contingencies, effectiveness scores were determined for each indicator. Effectiveness scores have a distinct meaning. Zero means that the shop is neither good nor bad, just meeting expected productivity. A positive score means the shop is exceeding expectations while a negative score means the shop is producing below expected productivity. Monthly feedback reports included performance on each indicator for that month, the effectiveness score associated with that level of performance for each indicator, overall effectiveness for the unit, and priority data for each indicator. The priority score is the amount of gain in effectiveness that is achieved from one unit improvement on the indicator. Indicators were presented ranked from highest to lowest priority, along with the priority scores, at monthly meetings of each work unit.

Assessment of Reliability and Validity

Steps were conducted to assure the reliability of contingencies. First, independent contingencies were developed for two shifts within each work unit. Contingencies were correlated .86 to .99. Effectiveness scores for each set were also compared. The average difference in contingencies across the two shifts was only 7.9%. Thus reliability across two different groups developing contingencies for the same work unit was high. The validity of the system was assessed using two different scenario procedures. One scenario had six supervisors rank team effectiveness based on scores on 13 indicators. A second scenario had these supervisors determine the effect of changes in indicators on team effectiveness. A high degree of agreement was obtained between supervisor judgments and ranks of team effectiveness obtained from actuarially applying the groups' contingencies. Pritchard et al. (1989) discuss the results of these assessments and argue for their support of the ProMES reliability and validity.

Method for Time Series Analysis

Bergh's (1993a) guidelines regarding time in model building suggest the need for: (a) a clear definition of the conceptual framework and measurement of time; (b) explicit incorporation of time induced effects in the model; (c) linkage between time context and analytical techniques; and (d) explicit testing of actual data conformance with statistical characteristics required for valid application of the analytical techniques. We follow Bergh by: (a) measuring time in terms of changes in performance from month to month for a number of work groups (i.e., this is a panel study); (b) explicitly including a lag between past performance and present performance (and by testing for further unspecified temporal effects in the error structure); (c) using panel analysis methods (generalized least squares regression) that treat time appropriately in a panel study; and (d) testing for the violation of the assumptions of independent errors in the work group cross section dimension (homoscedasticity) and in the temporal dimension (no autocorrelation), and by confirming absence of bias in the coefficients from multicolinearity through correlation analysis. In constructing our analysis, a theoretical model was first specified based on the foregoing hypotheses.

Specification of the Dependent Variable

The basic hypotheses of the present research relate to the idea that a work team's effectiveness, as measured by a set of relevant indicators, will improve if effort is allocated on a month-to-month basis to improve those indicators that can best benefit the team's overall effectiveness during that period. Testing these hypotheses must focus on determining if improvement on specific indicators has taken place. Improvement in effectiveness provides an appropriate measure because it is scaled in the same units across all indicators. Month-to-month improvement can be measured simply as the change in the level of the effectiveness score for an indicator over each month[4].

Specification of the Independent Variables in the Time Series Model

An intercept (**INTERCEP**) is kept in the equation to control for the possibility that effectiveness scores will change systematically over time even if there are no simple gain strategies, or more complex strategies based on feedback, goal setting, or incentive effects. The maximum potential gain in an indicator was used as a control variable (for regression to the mean and for any simple strategy related to low performance on a given indicator) by including a variable which measures the maximum gain in the effectiveness score possible from the effectiveness score obtained in the preceding month (**GAIN** = maximum effectiveness score (MAXEFF) - effectiveness in the preceding month (LEFF)). If this general effect is accounted for adequately and the effects of priority feedback, goal setting, or incentives are still found to be significant, then it is more likely that these latter effects are the result of the treatments.

For the first eight[5] months, no treatment was applied. Beginning with the ninth month[6] individuals in each unit were provided with feedback each month (FEEDBACK = 0 in each of months 1 to 8, and = 1 for each month thereafter) on their performance and were told what the priority score was for each indicator (PRI = the gain in terms of effectiveness from a single unit increase in performance on that indicator). Because most contingencies are nonlinear, priority scores for each indicator vary from month-to-month depending upon the level of performance for each indicator. Any change in effectiveness caused by awareness of a priority score for one period can only have an effect in the following period. Thus, the appropriate explanatory variable to use is the priority score lagged one period (i.e., month)(LPRI). Of course a relatively complicated concept such as the priority score could not have an effect until it was communicated explicitly to the teams. Thus, priority scores are hypothesized to only have effects following the beginning of the feedback period (Hypothesis 1). The variable used is **LPRIFEED** = LPRI * FEEDBACK computed for each month.

Hypothesis 2 states that each of the five work teams (shops) will have distinctly different experiences with respect to normal improvements in effectiveness over time without any treatments and will also have different responses to the priority score feedback. A set of zero-one dummy variables was used to test the possibility that work teams have different experiences regarding productivity improvements. Each observation was assigned a set of four 0 or 1 codes (**SHOP**$_i$ = 1 for the i^{th} shop, and 0 otherwise) with one SHOP$_i$ being omitted from explicit 0,1 coding to avoid perfect multicollinearity (Cohen & Cohen, 1975). In this case the Comm/Nav shop was coded with four 0's, thus making Comm/Nav the control shop against which the other four shops were compared. Interaction terms for the dummy variables with priority feedback were used to test the possibility that work teams responded differently to priority feedback (**FEEDSHP**$_i$ = FEEDBACK * SHOP$_i$) computed for each month.

Hypothesis 3 posits that formal goal setting affects effectiveness scores on productivity indicators beyond that afforded by potential gain, priority feedback, and interactions of shop with feedback. Since goals were set in terms of overall effectiveness relevant to the specific unit and the dependent variable is change in each specific indicator, it would be inappropriate to use goal level as the independent variable. Thus, we assessed the effects of the presence of explicit goal-setting activities on change in individual indicators of performance. Explicit goal-setting activities, which began in each unit in the 14th month, were accounted for with a zero-one dummy variable (**GOALSET** = 0 for each month 1 to 13, and = 1 for each month thereafter).

Hypothesis 4 concerns the effects of incentives on improvement in effectiveness on each indicator beyond that accounted for by potential gain, priority feedback, shop by feedback interactions, and goals. Incentives began in the 19^{th} month, were composed of additional time off with pay for all team members, and were offered related to overall effectiveness. Since incentives were the same for all units, they were accounted for using a zero-one dummy variable (**INCENTIV** = 0 for each month 1 to 18, and = 1 for each month 19 through 23). To confirm our expectations that no interactions of shop with goal setting or incentives will be necessary beyond accounting for shop by feedback interactions, the interactions of **SHOP**$_i$ with **GOALSET** and **INCENTIV** were included in the initial equation. In total there were 1,035 observations (23 months x 45 indicators across the five shops), one dependent and 20 independent variables.

<u>Specification of the Initial Model for Estimation</u>

The basic regression equation is of the following form:

$$DIF = INTERCEP + b_1 * GAIN + b_2 * LPRIFEED + b_3 * GOALSET + b_4 * INCENTIV + S_i\ b_i * SHOP_i + S_i\ b_i * FEEDSHP_i + S_i\ b_i * GOALSHP_i + S_i\ b_i * INCENSHP_i$$

RESULTS

For each of the effectiveness indicators in each shop, a simple plot of effectiveness scores over the 23 months was made. From these plots it was evident that there was a general tendency for effectiveness to rise over time, but the rates and patterns of increase varied considerably from indicator to indicator. If a unit's effectiveness score for an indicator does not vary over time because the effectiveness is already at the maximum for that indicator, then there is no chance of modifying the effectiveness and such indicators ought not to be included in the experimental frame. For example, the effectiveness score for "Number of Reportable Accidents" did not vary at all during the experimental period. It remained constant at 50 in each period, which is the score assigned to zero reportable accidents. Since improvement beyond this level is not possible, this indicator was omitted from the balance of the analysis. Six other indicators were omitted from this analysis for the same reason (cf. footnote [a] in Table 1). In our data, the only indicators that did not vary over time were those identified above which remained at their maximum for the entire test period. Dropping these seven indicators reduced the total number of observations to 874 (23 months x 38 indicators across the five shops). Using a difference variable as the independent variable and the lagged variable for priority scores reduces the total observations by the number of indicators in the first month to 836 total observations. The means, standard deviations, and bivariate correlations for the study variables described above are presented in Table 2.

Estimation Results

First, a simple ordinary least squares regression analysis which pools the data without recognition of its panel structure revealed (a) the hypothesis of first-order autocorrelation cannot be rejected (as indicated by the Durbin-Watson statistic of 2.28 for the model with 20 variables and 836 observations), (b) the hypothesis of homoscedasticity can be rejected (as indicated by an F statistic of 9.02 for White’s test), and (c) multicolinearity does not bias the coefficients as indicated by a maximum variance inflation factor of 1.85 for the continuous independent variables. Our use of the change in level of effectiveness from month-to-month rather than simply the level of effectiveness in each period should have removed the expected autocorrelation in levels of effectiveness from month to month for each indicator.

When the analysis is appropriately placed into a panel framework with an error structure that allows estimation of the time series effects and is estimated with generalized least squares instead of ordinary least squares, the model conforms with Bergh’s (1993a) guidelines and prescriptions for the appropriate treatment of data with both time series and cross section components. A generalized least squares estimation was conducted using a Fuller-Battese Variance Components Model (SAS Institute Inc., 1993). The GLS estimate of variance associated with the time dimension (the months)= 9.39, the variance associated with the indicators (cross section) = 97.90, and variance associated with error = 738.03. Thus, the variance is dominated by the cross sectional effects, indicating little variance due to time dimension beyond that accounted for by month-to-month change in effectiveness, the lagged effectiveness component of the GAIN score, and the lagged priority scores.

Initial estimation of the basic equation showed that, as expected, only the feedback regarding priority scores had differential effects among the various units. Interactions of $SHOP_i$ with goal setting and with incentives were trivial and not statistically significant. Since these interactions were not part of our theoretical model they were omitted. The resulting hypothesized model was:

$$DIF = INTERCEP + b_1 * GAIN + b_2 * LPRIFEED + b_3 * GOALSET + b_4 * INCENTIV + S_i\ b_i * SHOP_i + S_i\ b_i * FEEDSHP_i$$

The results along with parameter estimates for both the OLS and GLS analysis are reported in Table 3.

Interpretation of the Theoretical Model

The combination of the variables in the hypothesized model explained 21% of the variance in the month-to-month changes in effectiveness which is statistically significant ($F = 19.40$, $p < .0001$). The predicted values for changes in the level of effectiveness can be converted to predicted values for effectiveness itself (i.e., predicted value of the level in period 2 = level in period 1 + the predicted value for change in level from period 1 to period 2). Then the proportion of total variation in level of effectiveness that has been accounted for by the predicted levels can be calculated. In this case the combined variables explain almost 59% of the variation in the monthly level of effectiveness.

The GLS method again indicated that the majority of the variance accounted for is due to the cross section (97.99) with very little due to the time series effect (7.69). Comparing the OLS and GLS estimates, the parameter estimates from the GLS method are consistently higher for the significant effects than the OLS estimates.

Table 2. Means, Standard Deviations, and Bivariate Correlations of Study Variables

Variable	Number	Mean	Std Dev	DIF	VAR 1	VAR 2	VAR 3	VAR 4	VAR 5	VAR 6	VAR 7	VAR 8	VAR 9
DIF		1.848	32.381	1.000									
GAIN	VAR 1	38.833	43.857	0.361	1.000								
LPRIFEED	VAR 2	11.663	18.954	0.245	0.207	1.000							
GOALSET	VAR 3	0.455	0.498	-0.037	-0.461	0.081	1.000						
INCENT	VAR 4	0.227	0.419	-0.038	-0.290	0.025	0.594	1.000					
SHOP1	VAR 5	0.132	0.338	0.034	0.057	-0.090	0.000	0.000	1.000				
SHOP2	VAR 6	0.184	0.388	-0.021	0.145	0.249	0.000	0.000	-0.185	1.000			
SHOP3	VAR 7	0.158	0.365	0.003	0.007	-0.138	0.000	0.000	-0.169	-0.206	1.000		
SHOP4	VAR 8	0.237	0.425	0.002	-0.111	-0.040	0.000	0.000	-0.217	-0.265	-0.241	1.000	
FEEDSHP1	VAR 9	0.090	0.286	0.048	-0.122	-0.016	0.134	0.079	0.807	-0.149	-0.136	-0.175	1.000
FEEDSHP2	VAR 10	0.126	0.332	0.025	0.000	0.400	0.161	0.096	-0.148	0.798	-0.164	-0.211	-0.119
FEEDSHP3	VAR 11	0.108	0.310	0.014	-0.154	-0.062	0.148	0.088	-0.135	-0.165	0.802	-0.194	-0.109
FEEDSHP4	VAR 12	0.197	0.398	0.012	-0.127	0.145	0.211	0.126	-0.193	-0.236	-0.215	-0.276	-0.156
GOALSHP1	VAR 13	0.060	0.237	-0.013	-0.177	-0.057	0.276	0.164	0.648	-0.120	-0.109	-0.141	0.803
GOALSHP2	VAR 14	0.084	0.277	-0.019	-0.083	0.232	0.331	0.197	-0.118	0.636	-0.131	-0.168	-0.095
GOALSHP3	VAR 15	0.072	0.258	-0.010	-0.193	-0.102	0.305	0.181	-0.108	-0.132	0.642	-0.155	-0.087
GOALSHP4	VAR 16	0.108	0.310	-0.027	-0.191	-0.005	0.380	0.226	-0.135	-0.165	-0.150	0.623	-0.109
INCESHP1	VAR 17	0.030	0.170	-0.005	-0.124	-0.037	0.192	0.324	0.451	-0.083	-0.076	-0.098	0.559
INCESHP2	VAR 18	0.042	0.200	-0.003	-0.053	0.176	0.229	0.385	-0.081	0.440	-0.091	-0.116	-0.066
INCESHP3	VAR 19	0.036	0.186	-0.014	-0.135	-0.066	0.211	0.356	-0.075	-0.092	0.446	-0.107	-0.061
INCESHP4	VAR 20	0.054	0.226	-0.024	-0.131	-0.021	0.261	0.440	-0.093	-0.113	-0.103	0.428	-0.075

Table 2. Means, Standard Deviations, and Bivariate Correlations of Study Variables (Cont'd)

Variable	Number	VAR 10	VAR 11	VAR 12	VAR 13	VAR 14	VAR 15	VAR 16	VAR 17	VAR 18	VAR 19	VAR 20
FEEDSHP2	VAR 10	1.000										
FEEDSHP3	VAR 11	-0.132	1.000									
FEEDSHP4	VAR 12	-0.188	-0.172	1.000								
GOALSHP1	VAR 13	-0.096	-0.088	-0.125	1.000							
GOALSHP2	VAR 14	0.798	-0.105	-0.150	-0.076	1.000						
GOALSHP3	VAR 15	-0.105	0.801	-0.138	-0.070	-0.084	1.000					
GOALSHP4	VAR 16	-0.132	-0.121	-0.172	-0.088	-0.105	-0.097	1.000				
INCESHP1	VAR 17	-0.067	-0.061	-0.087	0.696	-0.053	-0.049	-0.061	1.000			
INCESHP2	VAR 18	0.552	-0.073	-0.104	-0.053	0.691	-0.058	-0.073	-0.037	1.000		
INCESHP3	VAR 19	-0.073	0.555	-0.096	-0.049	-0.058	0.694	-0.067	-0.034	-0.040	1.000	
INCESHP4	VAR 20	-0.090	-0.083	-0.118	-0.060	-0.072	-0.066	0.687	-0.042	-0.050	-0.046	1.000

Table 3. Theoretical Model Estimation Results with OLS and GLS Methods Analysis of Variance from the OLS Method

Source	DF	Sum of Squares	Mean Square	F Value	Prob>F
Model	12	193068.29	16089.02	19.40	0.0001
Error	823	682451.41	829.22		
Total	835	875519.71			

Root MSE	28.80	R-square	0.22
Dep Mean	1.85	Adj. R-sq.	0.21
C.V. 1557.17			

Variance Component Estimates from the GLS Fuller-Battese Method	
Root Mean Squared Error	27.29
Variance Component for Cross Section	97.99
Variance Component for Time Series	7.69
Variance Component for Error	733.66

Parameter Estimates

		OLS Method		GLS Method	
Variable	DF	Parameter Estimate	T for H_o: Parameter=0	Parameter Estimate	T for H_o: Parameter=0
INTERCEP	1	-19.64	-5.48	-24.18	-5.16
GAIN	1	0.37	12.40	0.47	14.92
LPRIFEED	1	0.15	2.24	0.30	4.40
GOALSET	1	5.08	1.74	9.17	2.84
INCENT	1	-0.70	-0.24	-0.19	-0.06
SHOP1	1	-15.87	-2.63	-20.60	-2.61
SHOP2	1	-17.55	-3.28	-20.80	-2.97
SHOP3	1	-11.27	-2.01	-14.70	-2.00
SHOP4	1	6.76	1.64	4.73	0.79
FEEDSHP1	1	29.52	4.54	31.92	5.07
FEEDSHP2	1	18.88	3.19	24.19	2.58
FEEDSHP3	1	22.48	3.81	24.19	4.22
FEEDSHP4	1	6.33	1.40	2.41	0.54

The intercept term for the overall theoretical equation is negative and significant. The GAIN parameter estimate was significant and positive. The provision of feedback regarding priority scores (LPRIFEED) had a positive impact on effectiveness beyond the impact of overall gain, confirming Hypothesis 1, and goal setting (GOALSET) had a further positive impact, confirming Hypothesis 3. The magnitudes of the coefficients show that the amount of potential gain (GAIN) produced a small increase (.37) in the monthly average effectiveness on indicators of performance. Since this parameter was estimated over the whole 23-month observation period, it would cumulate to an average increase across indicators of 8.51 effectiveness points. Providing feedback (in the form of priority scores) produced a small increase (.15) in the average size of effectiveness scores between periods. This increase would be cumulative over the last 15 months resulting in a total increase of 2.25 in the average effectiveness score across all shops. The coefficient of GOALSET shows that goal setting led to large increases in effectiveness of 5.08 units over the last 10 months, which cumulates to 50.80 units of increase in average effectiveness scores. Incentives were found to have no effect on productivity improvement in this data, even after controlling for the effects of goal setting and the interactions of $SHOP_i$ with feedback, disconfirming Hypothesis 4. The incentive effect was left in the equation so as not to confound our interpretation of the feedback and goal setting effects.

To understand the interactions we decomposed the equation by substituting the dummy codes for the various shops, and combining like terms to obtain separate equations for each shop (Cohen & Cohen, 1978). These equations are presented in Table 4.

Table 4. Decomposition of the Theoretical Equation by Work Group

THEORETICAL EQUATION:

$$\hat{Y} = .37\ GAIN + .15\ LPRIFEED + 5.08\ GOALSET + -.70\ INCENT + -15.87\ SHOP1 + -17.55\ SHOP2 + -11.27\ SHOP3 + 6.76\ SHOP4 + 29.52\ LPRIFEED*SHOP1 + 18.88\ LPRIFEED*SHOP2 + 22.48\ LPRIFEED*SHOP3 + 6.33\ LPRIFEED*SHOP4 + -19.64$$

COMM/NAV (CONTROL):

$$\hat{Y} = .37\ GAIN + .15\ LPRIFEED + 5.08\ GOALSET + -.70\ INCENT + -19.64$$

RECEIVING (SHOP 1):

$$\hat{Y} = .37\ GAIN + 29.67\ LPRIFEED + 5.08\ GOALSET + -.70\ INCENT + -35.51$$

PICKUP & DELIVERY (SHOP 2):

$$\hat{Y} = .37\ GAIN + 19.03\ LPRIFEED + 5.08\ GOALSET + -.70\ INCENT + -37.19$$

STORAGE & ISSUE (SHOP 3):

$$\hat{Y} = .37\ GAIN + 33.63\ LPRIFEED + 5.08\ GOALSET + -.70\ INCENT + -30.91$$

INSPECTION (SHOP 4):

$$\hat{Y} = .37\ GAIN + 6.48\ LPRIFEED + 5.08\ GOALSET + -.70\ INCENT + -12.88$$

The intercept terms for the separate shops are all negative. The finding that the shop coefficients for shops 1 (Receiving), 2 (Pickup & Delivery) and 3 (Storage & Issue) are all significant indicates that they are different from (more negative than) the control shop (Comm/Nav). The intercept for shop 4 (MS&D Inspection) was not significantly different from that of the Comm/Nav shop. This is represented in the magnitude of the potential gain (GAIN) in indicators early in the experimental period. During the baseline period the average maximum gain across indicators for MS&D Receiving, MS&D Pickup & Delivery, and MS&D Storage & Issue were 95.63, 80.88 and 82.60, respectively, while the average maximum gain during the same period for Comm/Nav and MS&D Inspection were 48.35 and 52.94, respectively.

It is interesting to note that the three shops with the lowest (most negative) intercepts (and greatest distance from their maximum productivity): (1) Receiving, (2) Pickup & Delivery, and (3) Storage & Issue all obtained significant positive parameters for the interaction of feedback with shop. This indicates that these shops had responses to feedback that were significantly more positive than the Comm/Nav shop (the three significant $FEEDSHP_i$ variables have coefficients of +29.52, +18.88, and +22.48). Thus, feedback of priority data appears to have a much greater positive effect on work units that initially perform more poorly. In fact, the positive effect of feedback almost offsets the initial poor performance so that all five shops are much more similar once feedback is provided. At the end of the feedback period (in month 13), the average maximum gain (GAIN) values for the initially lower performing groups (Receiving, Pickup & Delivery, and Storage & Issue) were much closer to the initially higher performing groups (Comm/Nav and Inspection) (10.40, 45.57 and 20.50 versus 38.73 and 12.78, respectively). Thus, the effect of priority feedback was to equalize poor performing teams to the level of the higher performing teams. The interaction of SHOP4 with feedback was not significant suggesting that the Inspections unit did not react differently from the Comm/Nav unit to the priority feedback.

DISCUSSION

This analysis provides a clear and valuable contribution beyond the previously published analyses of ProMES data. The inclusion of the intercept term and the inclusion of the GAIN variable in the regression equation provided control for time series effects, regression to the mean, and effects of simple strategies not requiring more sophisticated team strategies based on priority scores. Conducting the analysis at the indicator level allowed inferences to be made about work team performance. The fact that priority feedback provided significant effects after the intercept and the GAIN variables were controlled, supports our extension of NPI theory to team performance. These work teams used the priority information to guide productivity improvements (Hypothesis 1). Additionally, the response to feedback differed across teams confirms Hypothesis 2. The Comm/Nav and Inspection teams had better initial performance than the other teams. However, the strategy information provided by feedback allowed the other units to correct their strategies so as to reduce differences once feedback was provided. It is likely that providing specific feedback with priority data helped to clarify roles and expectations (Naylor, Pritchard, & Ilgen, 1980), possibly correcting management errors that led to prior poor performance among some teams.

Because priority feedback accounted only for an average 2.25 effectiveness point gain over the 15-month feedback period, one may be inclined to conclude that the limited contribution of priority data does not justify the cost of generating it. We argue otherwise. While specific priority feedback did not have a strong effect in the control group, specific priority feedback had substantial corrective effects among poorly performing teams. The strong interactions between priority feedback and shops indicates that feedback had an equalizing effect which brings those poor performing work teams up to the level of the better performing teams. In the present research, the fact that priority feedback information had strongest effects on poorly performing teams suggests that these teams had not recognized ineffective strategies. However, the ProMES procedures helped these teams to correct behaviors to conform to more effective strategies for the work group.

The present analysis shows that goal setting accounted for a cumulative average improvement of nearly 51 effectiveness points in indicators over the 10 months that goal setting was applied. This result can be interpreted as evidence that goal setting provided a strong motivational force behind the productivity improvements (Hypothesis 3). Because goal setting was implemented on top of feedback, we can compare only goal setting with feedback to feedback alone. We suggest that priority feedback (the prescriptive information indicating which indicators to improve) provided the strategy necessary to allow goal setting to have its effect. This is consistent with the goal-setting theory notion that feedback has little motivational effect apart from goal setting. It is interesting to note that there were no significant interactions between the $SHOP_i$ dummy codes and the goal-setting intervention. This may indicate that strategy differences across teams were eliminated by the provision of priority feedback and that teams were equally committed to their goals. This is consistent with the literature on task effects on goal setting which shows that task complexity moderates the effect of goal setting (Wood et al., 1987). Our evidence suggests that this may occur when groups have inadequate strategies for productivity improvement. However, if groups develop adequate strategies for complex tasks and/or for team coordination, the moderating effects of task complexity should be minimized. Our research was limited by the fact that we did not have specific information on task complexity and team member coordination. Future research on teams should investigate specifically the types of strategies that would be most effective given specific complexity and team coordination characteristics.

A strong commitment to goals may also explain why incentives had no effect on productivity (disconfirmation of Hypothesis 4), since incentives have been theorized to impact productivity through goal commitment (Wright, 1992). In addition to the possibility that goal commitment was high, it also appears that teams had attained near maximum effectiveness on most indicators by the time incentives were implemented (ceiling effect) (Pritchard et al., 1988, 1989). Additionally, it is possible that incentives may increase goal persistence over a longer period of time. To test this, researchers would have to continue incentives for experimental groups while discontinuing incentives for comparable control groups.

This research does not allow us to address the effects of goal setting without priority feedback. However, on the basis of past research (Locke & Latham, 1990; Wood & Locke, 1990), it seems reasonable to infer that the results of goal setting would not have been as strong without priority feedback as was observed in the present research. In combination with the effect size data presented by Pritchard et al. (1988), and recent research on goals and task and group strategies (Mitchell & Silver, 1990; Weingart, 1992; Weingart & Weldon, 1991; Weldon, Jehn, & Pradhan, 1991; Wood & Locke, 1990), we suggest that goal setting must be combined with a measurement and feedback system that allows for the development and continued monitoring of appropriate task strategies. Our findings suggest that, once appropriate strategies are developed within work groups, goal setting has consistent effects across those work groups.

SUMMARY AND CONCLUSIONS

This reanalysis demonstrates two important methodological points. First, while aggregate outcome measures such as those reported by Pritchard et al. (1988) are adequate for testing global outcomes of organizational interventions, disaggregating those measures at the indicator level and across time provides a clearer picture of the mechanisms by which interventions work. Second, conducting analyses in a manner that allows the control of time series effects and provides tests of the time series assumptions gives a clearer picture of the relative impact of motivational interventions implemented in a time series fashion. Analytical procedures must be designed to answer the specific questions proposed by theory and must take into account the nature of the data.

We argue that, even though these data have been published previously, this reanalysis both extends and adds critical information to organizational sciences. Specifically, we have applied the Naylor et al. (1980) notion of contingencies to work teams and tested this application using appropriate time series techniques with an appropriate data set. Work teams in this sample use both a simple gain strategy and a more specific strategy based on priorities developed for each indicator each month. Priority feedback benefited some teams more than others. The data suggest that teams which perform more poorly initially are benefited the most by priority feedback such that differences in performance across teams are reduced over time. Goal setting had substantial effects beyond feedback. The absence of any work team by goal setting interactions suggests that there were no work team-related moderators of the goal-setting effects. Thus, we conclude that the priority feedback provided work teams the necessary information to be able to benefit from the goal-setting intervention.

This analysis clearly demonstrates that teams afforded information in terms of priorities for specific indicators of productivity utilize that information to improve those specific work team products. While the research analyzed here did not directly assess work team strategies, we might infer that work teams developed strategies based on the priority information. While there are numerous methods to enhance the development and use of work team strategies, the ProMES method appears to be effective. More generally, we add to the growing literature identifying the importance of work team strategies to team performance.

REFERENCES

Argyris, C. & Schön, D. A. (1996). Organizational learning II: Theory, method and practice. Reading, MA: Addison-Wesley Publishing Company.

Bergh, D. D. (1993a). Watch the time carefully: The use and misuse of time effects in management research. Journal of Management, 19, 685-705.

Bergh, D. D. (1993b). Don't "waste" your time! The effects of time series errors in management research: The case of ownership concentration and research development spending. Journal of Management, 19, 897-914.

Cohen, J. & Cohen, P. (1975). Applied multiple regression/Correlation analysis for the behavioral sciences. Hillsdale, NJ: Lawrence Erlbaum Associates.

Earley, P. C., Connoly, T., & Ekegren, G. (1989). Goals, strategy development, and task performance: Some limits on the efficacy of goal setting. Journal of Applied Psychology, 74, 24-33.

Earley, P. C., Lee, C., & Hanson, L. A. (1990). Joint moderating effects of job experience and task component complexity: Relations among goal setting, task strategies, and performance. Journal of Organizational Behavior, 11, 3-15.

Klein, H. J. (1989). An integrated control theory model of work motivation. Academy of Management Review, 14, 150-172.

Klimoski, R., & Mohammed, S. (1994). Team mental model: Construct or metaphor? Journal of Management, 20, 403-437.

Locke, E. A., & Latham, G. P. (1990). Work motivation: The high performance cycle. In U. Kleinbeck, H. Quast, H. Thierry, & H. Hartmut (Eds.), Work motivation (pp 3-25). Hillsdale, NJ: Lawrence Erlbaum Associates.

Mitchell, T. R., & Silver, W. S. (1990). Individual and group goals when workers are interdependent: Effects on task strategies and performance. Journal of Applied Psychology, 75, 185-193.

Naylor, J. C., Pritchard, R. D., & Ilgen, D. R. (1980). A theory of behavior in organizations. New York: Academic Press.

Pritchard, R. D. (1990). Measuring and improving organizational productivity: A practical guide. New York: Praeger.

Pritchard, R. D. 1995. Productivity measurement and improvement: Organizational Case Studies. New York: Praeger.

Pritchard, R. D., Jones, S. D., Roth, P. L., Stuebing, K. K., & Ekeberg, S. E. 1988. Effects of feedback, goal setting, and incentives on organizational productivity. [Monograph] Journal of Applied Psychology, 73: 337-358.

Pritchard, R. D., Jones, S. D., Roth, P. L., Stuebing, K. K., & Ekeberg, S. E. (1989). The evaluation of an integrated approach to measuring organizational productivity. Personnel Psychology, 42, 69-115.

Riedel, J. A., Nebeker, D. M., & Cooper, B. L. (1988). The influence of monetary incentives on goal choice, goal commitment, and task performance. Organizational Behavior & Human Decision Processes, 42, 155-180.

SAS Institute Inc. (1993). SAS/ETS Users Guide (Version 6, Second Edition). Cary, NC: SAS Institute Inc.

Sawyer, J. (1990). Effects of risk and ambiguity on judgments of contingency relations and behavioral resource allocation decisions. Organizational Behavior and Human Decision Processes, 45, 85-110.

Sawyer, J. (1991). Hypothesis sampling, construction of adjustment: How are inferences about nonlinear monotonic contingencies developed? Organizational Behavior and Human Decision Processes, 49, 124-150.

Schmidt, F. L., & Hunter, J. E. (1993). Tacit knowledge, practical intelligence, general mental ability, and job knowledge. Current Directions in Psychological Science, 2, 8-9.

Schneider, S. C., & Angelmar, R. (1993). Cognition in organizational analysis: Who's minding the store? Organizational Studies, 14, 347-374.

Sternberg, R. J., & Wagner, R. K. (1993). The geocentric view of intelligence and job performance is wrong. Current Directions in Psychological Science, 2,1-5.

Sternberg, R. J., Wagner, R. K., & Okagaki, L. (1993). Practical intelligence: The nature and role of tacit knowledge in work and at school. In J. M. Puckett & H. W. Reese (Eds.), Mechanisms of everyday cognition, (pp 205-227). Hillsdale, NJ: Lawrence Erlbaum Associates.

Weingart, L. R. (1992). Impact of group goals, task component complexity, effort, and planning on group performance. Journal of Applied Psychology, 77, 682-693.

Weingart, L. R., & Weldon, E. 1991. Processes that mediate the relationship between a group goal and group member performance. Human Performance, 4, 33-54.

Weldon, E., Jehn, K. A., & Pradhan, P. (1991. Processes that mediate the relationship between a group goal and improved group performance. Journal of Personality and Social Psychology, 61, 555-569.

Wood, R. E., & Locke, E. A. (1990). Goal setting and strategy effects on complex tasks. In B. Staw & L. Cummings (Eds.), Research in organizational behavior (Vol. 12, pp. 93-109). Greenwich, CT: JAI Press.

Wood, R. E., Mento A. J., & Locke E. A. (1987). Task complexity as a moderator of goal effects: A meta-analysis. Journal of Applied Psychology, 72, 416-425.

Wright, P. M. 91992). An examination of the relationships among monetary incentives, goal level, goal commitment, and performance. Journal of Management, 18, 677-693.

FOOTNOTES

[1] The reader should not confuse “contingency” as used by Naylor et al. (1980) with the concept of contingent which reflects an interaction or moderating process. No interaction is implied by the concept contingency as defined by Naylor et al. (1980) , or as used in this paper.

[2] In the research analyzed here, all indicators are measured on ratio scales so that nonlinear transformations are meaningful.

[3] The reader is encouraged to read Pritchard et al. (1988) for details of the work unit sample and interventions, and Pritchard et al. (1989) for details on the development and validation of the ProMES System within this sample. More conceptual detail on the ProMES system can be found in Pritchard (1990), and others examples of applications of the system can be found in Pritchard (1995).

[4] In some circumstances the use of change or gain scores may create statistical problems. These difficulties have been well-documented in the literature. For example, Cohen and Cohen (1975) show how bias in using change scores can be removed with appropriate adjustments in both the variables and the model. In the present analysis these problems are reduced in two ways: (a) maximum gain possible is used as an explicit right-hand-side regressor to account for the potential bias resulting from large changes tending to follow low scores and low gains tending to follow high scores; and (b) time series autocorrelation in the residuals is explicitly treated separately.

[5] The baseline period prior to the application of any treatment is nine months in the Comm/Nav Shop. The length of the treatment periods for all shops is the same but the month number when each starts is one higher for the Comm/Nav shop for each treatment. To simplify the analysis, the first base period for the Communication/ Navigation shop was deleted leaving 23 monthly observations on each performance measure in each shop.

[6] Tenth month for the Comm/Nav Shop.