



University of Nebraska at Omaha
DigitalCommons@UNO

Student Work

7-1-1996

Consensus Decision Making of Quality Improvement Teams: A Descriptive Study

Ginger L. Riffel

University of Nebraska at Omaha

Follow this and additional works at: <https://digitalcommons.unomaha.edu/studentwork>

Recommended Citation

Riffel, Ginger L., "Consensus Decision Making of Quality Improvement Teams: A Descriptive Study" (1996). *Student Work*. 1593.
<https://digitalcommons.unomaha.edu/studentwork/1593>

This Thesis is brought to you for free and open access by DigitalCommons@UNO. It has been accepted for inclusion in Student Work by an authorized administrator of DigitalCommons@UNO. For more information, please contact unodigitalcommons@unomaha.edu.



**CONSENSUS DECISION MAKING OF
QUALITY IMPROVEMENT TEAMS:
A DESCRIPTIVE STUDY**

A Thesis

Presented to the
Department of Communication
and the
Faculty of the Graduate College
University of Nebraska

In Partial Fulfillment
of the Requirements for the Degree
Master of Arts
University of Nebraska at Omaha

by

Ginger L. Riffel

July, 1996

UMI Number: EP73533

All rights reserved

INFORMATION TO ALL USERS

The quality of this reproduction is dependent upon the quality of the copy submitted.

In the unlikely event that the author did not send a complete manuscript and there are missing pages, these will be noted. Also, if material had to be removed, a note will indicate the deletion.



UMI EP73533

Published by ProQuest LLC (2015). Copyright in the Dissertation held by the Author.

Microform Edition © ProQuest LLC.

All rights reserved. This work is protected against unauthorized copying under Title 17, United States Code



ProQuest LLC.
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106 - 1346

THESIS ACCEPTANCE

Acceptance for the faculty of the Graduate College,
University of Nebraska, in partial fulfillment of the
requirements for the degree Master of Arts,
University of Nebraska at Omaha

Committee

Name	Department
<i>Lyn K. Harland</i>	<i>Management</i>
<i>Hollis F. Gish</i>	<i>Communication</i>

Robert E. Carlson
Chairman *Communication*

7/11/96

Date

Abstract

This thesis reports on consensus decision making of quality improvement teams in the organizational setting. Specifically the study sought to determine whether a previously developed consensus instrument was a reliable and utilitarian measure of consensus. Analysis showed that the instrument was reliable, $\alpha = .9729$, but its usefulness remains in question. Additionally, other research questions addressed the relationship between external expert stakeholders' assessments of effectiveness and team members' assessments as measured by the instrument. Analyses showed only minimal relationship. As an unanticipated result in three tests, a negative relationship was found between one rater's rankings and a groups' consensus levels as compared to other groups' consensus levels. In tests of difference for sociodemographic variables, gender differences were found in the study, in that females consistently reported higher levels of consensus than their male peers. However, tests for the variables of age and title classification yielded no significant results.

Acknowledgements

I would like to thank my employer, University Hospital, for allowing me to complete my research using its Quality Improvement Teams. It is rewarding to work for an organization which truly values participation and employee development.

I would also like to thank the members of my committee. Dr. Lynn Harland's suggestions and scrutiny made this a better project. Dr. Hollis Glaser has helped broaden my perspective of small group communication and has made me more aware of issues of power in groups. A special thanks to Dr. Robert Carlson. His teaching and his guidance made this thesis the best learning experience of my graduate program.

To my coworkers and friends, thank you. Without your patience, I wouldn't be finishing this thesis. And to my family, words can't adequately express my thanks for your ongoing love, encouragement, and support.

TABLE OF CONTENTS

Abstract	iii
Acknowledgements	iv
List of Tables	vii
Chapter One	1
Introduction	3
Review of Literature	3
Group Decision Making	3
Quality Improvement/ Total Quality Management	10
Statement of Purpose	18
Chapter Two	20
Methodology	20
Setting and Subjects	20
Instrument	21
Procedure	22
External Stakeholder (Expert) Sample and Procedure	22
Statistical and Analytic Procedures	23
Chapter Three	24
Results	24
General Results	24
Results of Analysis for RQ 1	27
Results of Analysis for RQ 2	30
Results of Analysis for RQ 3	36
Chapter Four	38
Discussion	38

RQ 1	38
RQ 2	39
RQ 3	41
Chapter Five	43
Conclusions, Limitations, and Recommendations	43
Conclusions	43
Limitations	44
Recommendations	45
Recommendations for Quality Improvement Programs	45
Implications for Future Research	47
References	50
Appendix A	57
IRB Exemption	58
Appendix B	59
Cover Letter	60
Consensus Instrument	61
Appendix C	63
Comments	64

List of Tables

Table I: Frequencies of Response and Descriptive Statistics by Team	25
Table II: Frequencies of Response by Question	26
Table III: Rotated Factor Matrix: Consensus Instrument	28
Table IV: Consensus Instrument Reliability: Factor Means and Alphas	29
Table V: Team Effectiveness Ratings by Individual Raters	31
Table VI: One-way Analysis of Variance for Overall Consensus Instrument/ Expert Rater 3 Groupings	33
Table VII : One-way Analysis of Variance for Factor 1 (Group Decision)/ Expert Rater 3 Groupings	33
Table VIII : One-way Analysis of Variance for Factor 2 (Individual Opportunity)/ Expert Rater 3 Groupings	34
Table IX: One-way Analysis of Variance for Factor 3 (Relationships)/ Expert Rater 3 Groupings	34
Table X: One-way Analysis of Variance for Factor 4 (Individual Effectiveness)/ Expert Rater 3 Groupings	35
Table XI: One-way Analysis of Variance for Factor 5 (Decision Process)/ Expert Rater 3 Groupings	35
Table XII: T-Test for Mean Consensus Levels by Gender	37

CHAPTER ONE

Introduction

"Quality" is the buzzword of American business in the 1990s. Companies use the word freely in mission statements, strategic planning, and advertising campaigns. Businesses, the Armed Services, and nonprofit organizations implement Total Quality Management (TQM) or Quality Improvement Programs (QIP) to serve as forms of participative management, problem solving forums, and/or human resource development tools (Mohr & Mohr, 1983). Evolved from the Quality Circles of the 1970s and 1980s, quality teams work to address problems and to improve work processes; they are decision making groups.

While group decision making has long been of interest to scholars, many of the research findings may not be generalizable to groups in organizations. The vast majority of the research has been conducted in laboratory settings using self-contained, zero-history groups (Cragan & Wright, 1990). Subjects in these studies work on contrived decision making tasks for which there are right answers or preferred solutions. That is, the problems are not generally of the same nature as problems of "real world significance." According to Putnam and Stohl (1990), during the 1980s only 13% of communication studies were conducted in organizational or applied studies.

Limited research has been conducted on Quality Teams or their predecessors, Quality Circles. The majority of the literature on TQM is prescriptive in nature instead of being based on empirical evidence. Hundreds of articles are published annually in which proponents of Quality sell these programs as a panacea for any organizational or motivational problem (Park, 1991). However, relatively few research projects have been undertaken to gauge the effectiveness of these programs. Those research programs which

have measured effectiveness of Quality Teams have reported mixed or negative results (Park, 1991).

Quality teams are charged with consensus decision making. Consensus, viewed as the end state of members' feelings regarding group decisions, has several general meanings. First, the most common meaning of consensus is agreement with a groups' decisions (Davis, 1992; Hirokawa, 1980). Others, (Zalesnik & Moment, 1964; Klimoski & Karol, 1976) added dimensions of commitment and satisfaction with the group and its decisions as being central to understanding consensus.

Because quality teams are told that they must come to consensus regarding group decisions, the study of consensus in quality teams is appropriate and would provide some measure of group effectiveness. While the concept of consensus decision making has been widely studied, "the operationalization of consensus may be confusing and even contradictory across a number of studies" (DeStephen & Hirokawa, 1988). DeStephen and Hirokawa developed and tested a multidimensional measure of consensus in their 1988 study. They found their Consensus Instrument to be comprehensive and valid in the laboratory setting.

This investigation will attempt to replicate, in part, the study conducted by DeStephen and Hirokawa to determine if their instrument is a utilitarian measure of consensus in the organizational setting.

Review of Literature

Quality Improvement Teams are decision making groups, and Quality Improvement Processes have incorporated the results of group decision making research in their design. For example, a primary assumption of TQM and QIP is that groups make better decisions regarding complex issues than individuals. This assumption has been supported in some of the literature (Michaelson et al., 1989). Results of research on group process, consensus, and decision making methods are other areas which have been incorporated in the design of Quality Improvement Processes.

It is important, therefore, for this review of the literature to include both relevant group decision making research and the history of and previous research on Total Quality Management, or Quality Improvement Processes, as a framework for understanding the context of the present investigation.

Group Decision Making

Small group behavior has long been of interest to scholars, and the study of group decision-making has permeated a variety of academic disciplines (Gouran, 1984; Hirokawa & Johnson, 1989; McGrath & Kravitz, 1982). Scholars from fields such as psychology, economics, management, communication and political science claim the study of group decision-making within their domains. Much of the research on group decision-making has concerned the quality of decisions made by groups.

There has long been a debate as to whether individuals or groups make better decisions. In a 1982 review of related articles, Hill concluded that group performance was often inferior to that of the best member of the group. Miner (1984) also found that the best member, the individual in the group attaining the highest scores in a particular study, scored significantly above the group outcome.

On the other hand, Michaelsen, Watson, and Black (1989) reported that groups make better decisions. In their research, complicated tasks were developed to mirror those tasks which would be addressed in organizations. These researchers found that groups consistently outperform their best members. Using 222 groups, these researchers found that 100% of the groups outperformed the average single member score. Additionally, results indicated that 97% of the groups outperformed their best member.

Watson, Michaelson, and Sharp (1991) found that member competence leveled off during the course of their longitudinal study. That is, the most competent members in the initial stages of research were not the most competent members as the research continued. As groups gained experience, all members became "competent." This clearly supports the value of group decision-making, according to the researchers.

Other researchers have found no difference between the quality of decisions made by groups and individuals (Davis & Toseland, 1987; Yetton & Bottger, 1982). Davis and Toseland (1987) used graduate students in social work as subjects for their research. Subjects either worked in groups or worked individually using Social Judgment Analysis. As stated previously, no significant differences in the quality of decisions were found between groups and individual decision makers. Since the results of individual versus group decision-making research have varied, one would expect that this debate will continue. This area of group decision making research has important implications for Quality Improvement Processes, in that the superiority of group decision making is an underlying assumption of QIP.

Although some have claimed that the study of small groups is "the light that failed," research and theory building continue to grow (Baker, 1988). During the 1980s almost 100 published studies on small groups were produced by speech communication scholars (Cragan & Wright, 1990). Researchers have examined variables which affect

decision-making quality and various group processes in decision-making. Some of the research focuses on faulty decisions made by groups.

Turner, Pratkanis, Probasco, and Leve (1992) examined the effects of threat and cohesion on group effectiveness. In a test of Janis's theory of groupthink, the researchers found support for Janis's famous theory. Groupthink, as defined by Janis (1972), is the extreme concurrence sought by decision-making groups. Groupthink is likely to occur when a group experiences conditions like high cohesion, insulation from experts, limited data-gathering procedures, directive leadership and high stress coupled with low self-esteem and little hope in finding solutions which are better than ones suggested by an influential leader. These antecedent conditions result in symptoms of groupthink. Symptoms include stereotypes of outgroups, pressure on dissenters, illusions of invulnerability, self-censorship, and illusions of unanimity. It is hypothesized that groupthink results in poor decisions, and the concept has been applied to a number of well-known decisions, such as Ford's decision to market the Edsel and NASA's decision to launch the Challenger space shuttle (Turner et al. ,1992).

Other researchers have examined negative group experiences. Stohl and Schell (1991) examined negative and dysfunctional communication patterns in groups. They found that a "farrago," or organizational bully, often results in confusion as to responsibilities in a group, the group task, and group decisions. Stohl and Schell maintain that this role arises from and in interaction of the group. The communicative profile of the farrago includes creating crises, using any information, half-truths and innuendos, in a way that is useful, trivializing others' issues, and infiltrating communication networks and issues beyond his or her scope. In this manner, the farrago plays an expanded role in a group and ultimately affects the group's decisions.

Conflict in and satisfaction with groups have also been the objects of some research on group decision-making. Wall, Galanes, and Love (1987) explored the interrelationships of conflict, conflict management, member satisfaction, and quality of outcome in small, task-oriented groups. They suggested a curvilinear relationship exists between the number of conflict episodes in a group and the subsequent quality of the decision. They also found that integrative conflict-management strategies were associated with higher quality decisions than were distributive strategies. Satisfaction was found to be negatively related to conflict.

In a similar study, Wall and Nolan (1986) focused on individual satisfaction, type and amount of conflict, conflict resolution, perceived inequity, and quality of outcome. The results of their study indicated that greater inequity was associated with people conflicts which were managed by avoidance. Greater satisfaction was associated with task conflicts which were managed integratively. Inequity was negatively related to satisfaction, positively related to conflict, and was not found to be related to outcome quality.

In an exploratory study on conflict and decision-making, Pendell (1990) uncovered six categories of deviant behavior. Those group members who missed meetings or did not participate fully were labeled in the category, *lack of participation*. Another role was the *personality deviant*. *Testing others' opinions and solutions*, or opinion deviance, was another category uncovered in this study. The three other categories were *incompatible, leadership, and coalition* deviance (Pendell, 1990, pg. 400). Results of this study indicated that not all deviant behaviors lead to conflict situations; however, *opinion* deviance was related to conflict initiation.

Other researchers have considered different variables in group decision-making. Baker (1988) found that age, participation in the first minute, gender, and size all had

effects on total participation in groups. These findings have implications for research on decision-making groups because participation hierarchies develop as stable entities very early in group discussion and have been shown to influence group decisions.

The question of how men and women behave as members of small groups has been the subject of study and controversy for many years. For example, Sever (1991) found that dominance by men in groups led to high rewards and more positive ratings on affective dimensions than did dominance by women. This result suggests an interaction between sex and dominance. Sever's research results were similar to early studies by Strodbeck et al. (as cited in Mabry, 1985). However, not all of the research concludes that men emerge as role specialists in small groups. Mabry (1985) found that groups composed predominately of men were significantly lower in dominance acts under unstructured task conditions than were all-male groups assigned a highly structured task. Mabry suggests that small group interaction may not be substantively affected by gender composition. While men and women do interact somewhat differently as group members, the conditions under which the differences are most likely to occur have not been clearly identified. According to Mabry, attaining some gender mix in task groups may be sufficient. It would seem, however, that gender remains an important variable in the study of small group decision-making.

Various group processes which affect decision-making have also been studied extensively. And the attempts to improve group problem-solving have concentrated mostly on altering group processes. Hackman and Morris (1975) found that training in group dynamics would have a positive effect on group problem-solving.

Other research has centered on improving the member resources of the group through training in effective search and evaluation routines before the group starts the problem-solving process. Bottger and Yetton (1987, pp. 656-657) asserted that "group

problem-solving performance might be improved, at least, as efficiently as, if not more effectively than process consultation, by an intervention that improves member task contribution." However, their research design was flawed, according to Ganster, Williams, and Poppler (1991). While Bottger and Yetton's training program addressed general decision-making, Ganster et al. found that task-specific training may be more appropriate. These researchers argued that organizations would be wiser to concentrate their training resources in improving the technical knowledge of their decision-makers rather than improving the ways in which they use that knowledge.

Groups spend a significant amount of time discussing problems and generating solutions for those problems. Van de Ven and Delbecq (1974), in a landmark study, focused upon three alternative methods for group decision-making: interacting, nominal and delphi processes. The criteria chosen to measure the effectiveness of the three methods were the number of unique ideas developed and the satisfaction of groups with the decision-making process. On an applied fact-finding problem with no known solution, the researchers found nominal group technique and the delphi method to be equally effective, and both were superior to interacting groups.

However, interacting groups, or group brainstorming, remains a popular process in groups, despite the evidence that individuals produce fewer ideas in interactive brainstorming than when brainstorming alone (Van de Ven & Delbecq, 1974). Paulus, Dzindolet, Poletes and Camancho (1993) attribute this popularity to the illusion of group productivity in brainstorming. They found that most individuals believe they would generate more ideas in a group than they would individually. Two factors seem to contribute to the illusion of group productivity. First, results indicate that there is opportunity for social comparison in group brainstorming. When people brainstorm in groups, there is a focus on the group instead of the individual. Generally, individuals

compare their performance as being similar to others. Further, it also appears that individuals take credit for "a disproportionate amount of the brainstorming activity in groups" (Paulus & Dzindolet, 1993, pg. 585). Paulus and his associates have identified certain factors which may solve the riddle of productivity loss in brainstorming. However, they suggest future research on blocking, evaluation apprehension, free riding, social matching, and the perception of productivity.

Consensus is another aspect of small group decision making which has received much attention from researchers. However, consensus has not been operationalized in the same way in much of the research (DeStephen & Hirokawa, 1988). Schwenk and Cosier (1992) described consensus as an aid to or means of strategic decision making. On the other hand, when consensus is viewed as an end state of decision making, other general meanings emerge.

The most common meaning of consensus, viewed as an end state, is agreement with the groups' decision (Hirokawa, 1980). Davis (1992, pg. 3) also defined consensus in a general way, denoting consensus as "mere collective agreement on a choice, judgment, opinion or the like and (implying) no particular process, rule, or criterion." Zalesnik and Moment (1964) added the element of commitment to the group and its decision as another dimension of consensus. Other researchers have used member satisfaction with the group and with their individual contributions as indications of consensus (Burgoon, 1977; Hrycenko & Minton, 1974).

Based on earlier definitions of and research on consensus, DeStephen and Hirokawa (1988) developed a multidimensional instrument to measure consensus in decision making groups. They operationalized consensus as being comprised of five dimensions: feelings regarding the group decision, feelings regarding the decision process, feelings regarding group member relationships, feelings regarding individual effectiveness,

and feelings regarding opportunity to participate. The results of their research indicated that their instrument was a "statistically sound and conceptually defensible measure of small group consensus (DeStephen & Hirokawa, 1988)."

Quality Improvement/Total Quality Management

The forms and extent of employee participation in the United States have become diverse and numerous (Russell, 1988). Employees now participate in ownership or profits of organizations, some organizations are controlled and owned by employees, and employees participate in decision-making activities. One of the most popular forms of participation or involvement is commonly referred to as quality circles or quality teams (Vandervelde, 1979).

In the early 1970s, quality circles (QCs) became one of the most popular management interventions in this country. In that year, Lockheed Missile and Space Company introduced quality circles in its production areas. Although QCs are considered a Japanese innovation, two Americans, Edward Deming and J. M. Juran, are credited with QC development in post-World War II Japan (Russell, 1988). Originally, both Deming and Juran attempted to introduce their ideas in this country. However, American businesses flourished after World War II; companies didn't see the value in employee involvement even though social scientists had long advocated giving employees a direct voice in the workplace. Lewin et al., 1939, McGregor, 1957, and Argyris, 1960, had all suggested the practice of participation (Buch & Spangler, 1990). However, in the boom years following the war, there weren't many instances of participation, and Deming and Juran took their ideas to Japan.

Environmental changes during the 1970s and 1980s prompted many U. S. organizations to explore and experiment with work innovations such as employee

participation (Park, 1991). Foreign competition, decreasing productivity and quality, the general decline in the American economy, lower morale of employees, and the changing political climate all served to force managers in this country to change their styles of management. Lockheed had been enormously successful in implementing the Japanese QCs; they documented a \$2,844,000 savings during the first two years of their program (Smeltzer & Kedia, 1985). Following the success of Lockheed, other organizations began experimenting with QCs.

Lawler and Mohrmann (1987) have cited three reasons for the growing popularity of Quality Teams: American companies are trying to replicate the industrial success of Japan; "fadism" has always been prevalent in American business; and only limited structural changes are introduced as opposed to more extensive participative mechanisms. Therefore, the basic structure of the organization remains intact.

As the popularity of QCs has grown, organizations other than industrial companies have implemented quality initiatives and have changed the form of Quality Circles to some extent. Today, thousands of quality programs of various types exist in this country (Townsend & Gebhardt, 1990). For purposes of this discussion, the terms "Quality Circle (QC)", "Quality Teams" and "Total Quality Management (TQM)" will be used.

One can attribute this popularity of TQM to the expectations of American management; their interest in and support of Quality Teams is based on four assumptions (Ferris & Wagner, 1985). The first assumption is that workers know best what needs to be improved in the workplace and in the work they do. A second assumption is that groups outperform individuals in the identification and solution of organizational problems. The third assumption is that participation enhances productivity. Finally, the model of TQM assumes that American workers desire participation in the workplace. Some of these assumptions have been tested in the research.

Landon and Mouton (1986) found, through a survey of unionized firms, that product quality and productivity had increased after the firms had implemented quality initiatives. Additionally, they found that employees developed increased skills, increased their self-esteem, and developed an ability to make suggestions and implement changes.

Other researchers have concurred with the finding that participation in TQM increases productivity. Buch and Spangler (1990) discovered that in the year following involvement, Quality Team employees received significantly higher performance ratings and were promoted more frequently than nonmembers. While these researchers maintained that the developmental properties of the TQM process were believed to be responsible for the findings, Buch and Spangler also admitted that visibility, positive evaluation bias, and anticipatory socialization were also possible explanations for the results. These researchers cautioned that employee development should not be expected by merely implementing teams. Instead they suggested that employee development must be clearly articulated as a program objective, and that developmental activities be included in the program. Steel and Shane (1984) also showed positive results in increased productivity in their research on quality circles.

Thompson (1982) made sweeping statements about the results of quality circles. This scholar stated that QCs improve morale, increase a sense of loyalty to the organization, and foster a sense of teamwork. He also claimed they improve productivity, and they improve the quality of the product or service. QCs were also credited with reducing grievances, accidents, rework, absenteeism, and tardiness.

In a 1986 study, Stohl suggested that the emergent networks of members of quality teams have consequences for cognitive processing, member's attitudes, perceptions of the communicative climate, and the overall effectiveness of the organization. Her results indicated that TQM programs provide structures and encourage communication

that cut across the functional and hierarchical divisions of an organization. They may also change the basic culture of an organization; in short, they can be a powerful management tool. Others have agreed with Stohl's assessment. Schonberger (1994) cited culture change as being an important outgrowth of TQM. However, not all scholars agree with this assessment. Tuckman (1994, p. 729) contended that the TQM process may "paradoxically lend itself to the bureaucratization process."

Another assumption upon which TQM is based is that American workers desire participation. Fenwick and Olson (1986) found that high levels of aspiration for participation in decision-making exists in the American workforce, particularly among unionized employees and younger and female employees. Liverpool (1990) discovered that QT members perceived they had some say in work-related decisions; however, neither QT members nor nonmembers expressed a desire to have more than some say on most policy-oriented decisions.

On the other hand, some researchers have found evidence which contradicts the assumption that workers desire participation. Hackman and Oldham (1980) discovered that some American workers resist taking on enriched tasks that involve increased job autonomy and accountability. Other studies have found that people unaccustomed to participating in work place decision-making may attempt to avoid doing so, and that people's need for challenges and personal growth decline through periods of deprivation. In other words, in an authoritarian environment, people become reluctant to participate or voice opinions. (Ferris & Wagner, 1985).

While most of these results look promising to those involved in quality programs, not all results and reviews of quality circle studies are positive. Kanter (1982) suggested that Quality Teams don't provide workers with any actual control, but merely with the illusion of control by offering them a chance to provide input that the organization

subsequently ignores. She suggested that workers may view participation in TQM as inauthentic.

Others have been even more critical. Wendt (1994, pg. 5) claimed, from the results of a case study at a major university, that the TQM philosophy and processes are potentially dysfunctional in two aspects. He concluded that the TQM "hegemony" has the potential to reify linear and dualistic thinking, as well as having the potential to reprioritize traditional values of higher education. He stated that efficiency, cost effectiveness, and productivity may replace experimentation, the inherent value of ideas, and critical/creative pedagogy. The relative merits of TQM seem to remain in question.

Besides testing assumptions about TQM, some of the effects tested in the research are absenteeism, job satisfaction, organizational commitment, improved morale, increased productivity, cost-savings, employee attrition, promotion, and increased member suggestions. Donovan and Jury (1983) found improvement on nine of nine indices of attitudes in their study of hospital employees, as well as a 46% reduction in costs over two years. Tortorich et al. (1981) examined production quantity, quality and rework costs in their study of production personnel involved in TQM. They found a significant increase in three of four productivity criteria for Quality Teams.

While some researchers have found positive results, other observers are more neutral, holding that TQM can lead to success or failure, depending on how each particular program is implemented (Steel & Shane, 1984). Steel et al. (1985) investigated the outcomes of quality programs started in two organizations, a hospital and a facilities maintenance organization, located at a U. S. Army installation. The researchers found positive results on 7 of 20 criteria in the data from the maintenance sample, including reduced absenteeism, increased promotion, and increased productivity. On the other

hand, data from the hospital sample showed no positive effects. In fact, the hospital data revealed significant negative trends in attitude change among the members.

These researchers concluded that there was a significant difference in the manner in which the two programs were implemented which may have accounted for the dramatic differences in program outcomes. A higher level of management support was present in the maintenance sample. Additionally, more systematic and complete member training appeared to yield positive results in that sample.

Stohl and Jennings found mixed results in their quality study. They discovered that workers who voluntarily join Quality Teams have less job satisfaction but more organizational loyalty than those who don't join (as cited in Cragan & Wright, 1990). Finally, Gladstein (1984) reported that group satisfaction and effectiveness are linked to communication openness and supportiveness within the group and to boundary spanning communication outside the group.

In addition to producing mixed results, the research conducted on quality improvement has been faulted on other grounds. Pretest-posttest designs have been used quite often to evaluate programs; these designs have been faulted extensively because they fail to provide controls for Hawthorne or novelty effects (Steel & Shane, 1986).

Steel and Shane (1984) also cite small sample sizes and insufficient statistical power in those studies which aggregate data and employ groups as the unit of analysis. While longitudinal studies are desirable, some research efforts have used military settings and have experienced excessive subject mortality because of high turnover rates (Mento et al., 1984).

Abbott (1987) suggested that a number of empirical studies have pointed to the failure of quality circles to achieve their desired results. He suggested that scholars should

recognize the complexity of TQM and focus research and intervention efforts on individual, team, and organizational levels.

Putnam and Stohl (1990) have also criticized the research on TQM. They stated that research on Quality Teams in organizations has controlled the intensity and range of naturally occurring behaviors to test for group satisfaction and productivity. They concluded that simply moving small group studies into the field would not capture the dynamics of real-world groups.

Two studies have shown how decisions, proposals and arguments emanate from a quality group's external context. Geist and Chandler (1984) and Sabourin and Geist (1990) conducted discourse analyses on groups' discussions. Both of these studies revealed that the organizational context relates to the content and nature of the group deliberations.

Another study which considered the external environment was a 1990 study by Ancona and Caldwell. These researchers undertook an exploratory/descriptive study of new product teams within an organization. External judges gauged the effectiveness of teams in terms of quality and communication. Those groups who had frequent contact with external stakeholders were judged most effective. These researchers proposed that group performance would be enhanced if the amount of external activity increases as resource dependence increases. They also proposed that group performance would be enhanced if there were high levels of "scout" and "ambassador" behaviors. These behaviors include gathering data of various types and keeping others' informed. Finally, they proposed that groups would become more effective if they shifted emphasis between internal and external activities.

Putnam and Stohl (1990) have called for studying bona fide groups, and have examined quality teams as "bona fide" groups. They maintain that the study of groups in

their contexts involves much more than adopting an open systems view. They believe that bona fide groups are both the product and the producer of interactions that evolve and that the boundaries of bona fide groups are fluid. They suggest that future studies should assess the perceptions of group interactions with external stakeholders. Their view of bona fide groups and their suggestions offer an exciting framework for research. They state that adopting this view can revitalize research and inspire the imagination.

Statement of Purpose

Quality Improvement Teams are bona fide groups working in a larger organizational context. They are charged with coming to consensus, or agreement, on complex issues/problems for which there are no "right" answers. Most often the research conducted using these teams has concentrated on affective results and organizational benefits.

While there is a sizable body of research on group decision making, including consensus decision making, much of that research has been carried out in laboratory settings using zero-history groups. Therefore, the results of this research may not be readily generalizable to groups working on problems of real world significance.

Understanding consensus is central to understanding the decision making of Quality Improvement Teams. However, the concept of consensus has been operationalized in various manners in the decision making literature, as discussed in the review of literature. DeStephen and Hirokawa (1988) found that five dimensions appear intrinsic: agreement with the group decision, commitment to the decision, satisfaction with the decision, satisfaction with individual participation in the decision making, and satisfaction with the group decision making process.

The purpose of this study is to examine the consensus decision making of Quality Teams within the organizational setting. In essence, the project has two primary goals: to determine if the particular measure of consensus developed by DeStephen and Hirokawa proves a utilitarian measure of consensus with quality teams from an academic health sciences center, and to investigate the relationship between expert external stakeholders' assessments of effectiveness and team ratings of effectiveness as measured by the Consensus Instrument.

Specifically, three research questions will be addressed:

RQ1: Can the reported reliability and the factor structure of the DeStephen and Hirokawa Consensus Instrument be replicated?

The DeStephen and Hirokawa instrument has been tested only in the laboratory setting with university students. To determine if the instrument is a utilitarian measure of consensus, it is important to attempt to replicate both the reported reliability and the factor structure of the Consensus Instrument.

RQ2: Is there a relationship between expert rankings of team effectiveness and team ratings of effectiveness as measured by the Consensus Instrument?

Because Quality Improvement Teams operate within the organization and make decisions which affect the organization's processes, their assessments of effectiveness should be compared with external stakeholders' assessments of team effectiveness.

RQ3: Regardless of task, are there differences in the levels of consensus based on sociodemographic variable of age, gender, or title classification?

While this research question is not central to the primary goals of this thesis, it may yield information which has important implications for the formation of future Quality Improvement Teams. There is some evidence that age, sex, and status may affect satisfaction with decision making processes (Tang et al., 1989). For this reason, the research question is included in an attempt to gain further understanding of these variables' impact in team decision making.

CHAPTER TWO

Methodology

Setting and Subjects

The project site for this descriptive study was University Hospital, which is part of the University of Nebraska Medical Center. University Hospital is a 350 bed quaternary and tertiary care center

In 1992, University Hospital began implementation of Quality Improvement Teams as part of its "Distinction Through Quality" initiative. Quality Improvement projects are chosen by a Quality Council. The Quality Council, which consists primarily of administrative personnel of the hospital and clinics, chooses projects according to a certain set of criteria. All projects must contribute to the strategy; they must be issues which involve several functional areas; and they must present the potential for savings of resources or for increased efficiencies. Team members are invited to participate in particular projects because of familiarity with the issue or because their area is impacted by the issue.

Team members participate in a one-day, initial training session before they begin working as a team. They receive training in the principles of quality improvement. The principles are given to guide teams' work and include focusing on the customer, preventing problems, making decisions based on facts, focusing on improvement -- not blame, and continuously improving the system. Teams also learn about group stages and team building. Finally, these sessions include training in various methods of decision making and the use of quality tools, such as flowcharting, pareto analysis, and affinity diagramming. Each team is also assigned a facilitator and a management engineer. The trained facilitator aids in the group process, monitors group relationships, and provides

"just-in-time" training. The management engineer helps with data collection methods and analysis.

Teams are given a charter of their particular issue. First they have the opportunity to revise the charge to the team as they deem appropriate. Teams are given approximately 12 to 15 weeks to gather data, problem-solve, and finalize recommendations. At the end of their projects, teams make a formal presentation of their recommendations to the Quality Council.

This study was carried out through participation of 20 quality improvement teams. These teams were the total number of teams which had completed their work and presented their recommendations as of October, 1995. Quality Improvement Teams generally have 8 to 14 members. All members of all teams were included as subjects for this study.

Instrument

The 21-item consensus instrument, (DeStephen & Hirokawa, 1988), was used in this study. This instrument had been tested by DeStephen and Hirokawa (1988) in a pilot study and was found to be statistically and conceptually valid. The instrument was presented in Likert fashion and contained scales for each of five dimensions of consensus: feelings regarding the group decision, feelings regarding the decision process, feelings regarding group member relations, feelings regarding individual effectiveness, and feelings regarding individual opportunity to participate. Sociodemographic questions were added to the instrument to capture the variables of age, gender, and title classification. Finally, an area for comments was included at the end of the questionnaire.

Procedure

The Hospital Quality Officer at the subject institution approved this project. After approval by the student's committee and the University Internal Review Board (See Appendix A), a cover letter and the instrument were sent to all subjects (See Appendix B). The cover letter and instrument were printed on 11x 17 inch paper as one document. Subjects were asked to complete the instrument and to return it in a pre-addressed return envelope. A two-week time frame for return was suggested. A follow-up postcard was prepared to send to all subjects as a reminder. However, because the response rate was over 65% at the end of 10 days, it was not mailed.

External Stakeholder (Expert) Sample and Procedures

Three external stakeholders were chosen subjectively to assess the effectiveness of the decisions of the 20 quality teams. They were chosen because of their familiarity with and roles regarding the "Distinction Through Quality" initiative. Raters were either members of the Quality Council or staff of the Continuous Quality Improvement Implementation department which trains teams in the quality processes and coordinates the "DTQ" initiative.

Arrangements were made to meet with each expert rater individually. Raters were thanked for participating in the thesis study and were given the following instructions, "Because of your familiarity with the Quality Improvement Teams which have concluded their work, I'm asking that you rate them, in terms of the effectiveness of their decisions. Please keep in mind that I'm asking about the team effectiveness at the time they presented their findings to the Quality Council." Raters were given 20 4" x 6" cards with the name of a Quality Team printed on each one. Four rating cards were placed on the table with the following labels: Highest Level of Effectiveness, Second Highest Level of Effectiveness, Third Highest Level of Effectiveness, and Fourth Highest Level of

Effectiveness. Raters were then given the instructions to arrange the twenty cards in the four categories, putting only five cards in each category. Raters were also given a typewritten sheet with all twenty teams to serve as a guide.

Statistical and Analytic Procedures

The DeStephen and Hirokawa study employed Cronbach's alpha (1951) to verify the reliability of the scale. The resulting alpha was .8906. This study also employed Cronbach's alpha to answer, in part, the first research question, "Can the reported reliability and the factor structure of the DeStephen and Hirokawa instrument be replicated?" A principal-components factor analysis using varimax rotation was also used in the analysis.

In order to answer the other two research questions, descriptive statistics and tests of difference were used. Code numbers were randomly assigned in order to identify the individual teams. The third research question asks, in part, if there are differences in the levels of consensus based on the sociodemographic variable of gender. To answer this question, t-tests were performed. A series of one-way analyses of variance (ANOVAs) with follow-up Student Newman-Keuls procedures were used to test for differences in consensus levels based on the variables of age and title classification. ANOVAs were also employed to answer the question of whether there was a relationship between expert rankings of effectiveness and group mean levels of consensus.

CHAPTER THREE

Results

General Results

Of 136 consensus instruments sent to quality team members, 95 were completed and returned, which represents an overall response rate of 70%. Of the 95 respondents, 40 were male, and 55 were female.

In terms of other sociodemographic variables, respondents were asked to identify themselves by age and title classification. Of the 95 team members responding, 6 were between the ages of 20 and 29, 33 were between the ages of 30 and 39, 40 were between the ages of 40 and 49, and 16 were between the ages of 50 and 59.

All employee positions at University Hospital are categorized in one of three title classifications. Twenty-one respondents identified themselves as faculty/administrative; 69 identified themselves as managerial/professional; and 5 identified themselves as clerical/support personnel.

An analysis of frequencies indicated that responses were obtained for each of the 20 teams included in the study. Response rates ranged from 33% to 100% of team members responding. Descriptive statistics indicated overall mean scores for all 21 items by team ranging from 2.81 to 4.60. (Potential ranges would be from 1 to 5.) Frequencies and Descriptive statistics for each team are presented in Table I.

Frequencies of response by question were obtained for the 21 items on the consensus instrument. The results of this analysis are reported in Table II.

Table I: Frequencies of Response and Descriptive Statistics by Team

Team I.D.	# Sent	# Responding	Overall Mean	S.D.
1	5	4	4.30	.47
2	9	8	4.18	.47
3	4	2	2.81	1.75
4	7	4	3.96	.56
5	4	3	4.57	.30
6	8	7	4.07	.35
7	9	5	3.11	1.42
8	5	2	4.02	.57
9	9	5	4.49	.22
10	8	6	4.02	.27
11	7	3	3.97	.50
12	6	6	3.83	.22
13	6	5	3.91	.26
14	8	5	3.74	1.24
15	5	5	4.57	.24
16	6	2	4.00	.61
17	7	6	4.60	.26
18	6	5	3.71	1.33
19	8	5	4.04	.61
20	9	7	3.46	.83
TOTALS	136	95		

Table II**Frequencies of Response by Question**

Question #	Strongly Disagree (1)	Disagree (2)	Neutral (3)	Agree (4)	Strongly Agree (5)
1	3	3	7	45	57
2	4	2	6	42	41
3	4	2	4	34	51
4	1	6	7	39	42
5+	2	6	8	33	44
6	2	7	17	39	30
7++	5	3	12	37	37
8	2	6	7	45	35
9	2	9	10	43	31
10	3	4	11	40	37
11	3	4	6	36	46
12	7	4	16	41	27
13	0	11	34	34	16
14	1	7	8	50	29
15	1	11	17	49	17
16	2	8	9	59	17
17	3	25	31	29	7
18	4	3	3	49	36
19	0	5	10	64	16
20	1	5	10	58	21
21	3	3	8	50	31

N = 95

+ (n = 93)

++ (n = 94)

Finally, an area for comments was included at the end of the questionnaire. The complete comments are presented in Appendix C. In general, 11 respondents commented positively about the team and/or team experience. Twelve respondents commented that the recommendations made by their teams had not been implemented or had not been fully implemented. Eight respondents commented negatively about some aspects of team methods or efficiency.

Results of Analysis for RQ1

RQ1: Can the reported reliability and factor structure of the DeStephen and Hirokawa Consensus Instrument be replicated?

This study repeated the analysis measures of the DeStephen and Hirokawa (1988) study to verify the reliability of the scale (Cronbach, 1951). The reliability analysis (Cronbach's alpha) yielded the resulting alpha = .9729, which was higher than the alpha of .8906 reported in the DeStephen and Hirokawa (1988) study.

A principal-components factor analysis using varimax rotation, as was used by DeStephen and Hirokawa, demonstrated that two distinct factors accounted for a cumulative total of 73% of the variance. (See Table III).

While the factors did not cluster as they did in the study by DeStephen and Hirokawa (1988), the factors support the conceptual development of the instrument in that Factor 1 (Items 1 through 12) consisted of items assessing team members' feeling of agreement with the team decision, the team process, and team relationships. Factor 2 (Items 14, 15, 16, 17, 19, 20, and 21) reflected members' feeling of individual effectiveness and individual opportunity to participate. Two items did not clearly load on either factor. Item 13 read, "We were a closely knit team." Item 18 read as follows: "During team meetings, I got to participate whenever I wanted to."

Table III
Rotated Factor Matrix: Consensus Instrument

Item	Factor 1	Factor 2
1	.84146*	.30818
2	.87805*	.31664
3	.87556*	.34322
4	.87639*	.22687
5	.77124*	.32987
6	.79009*	.19716
7	.83135*	.17306
8	.82090*	.27344
9	.75872*	.33227
10	.73304*	.43977
11	.75115*	.42860
12	.67542*	.37924
13	.55484	.47504
14	.44544	.78802*
15	.32057	.83701*
16	.35148	.83591*
17	-.02002	.63752*
18	.59983	.53892
19	.58280	.60068*
20	.55221	.67408*
21	.58218	.71279*

* Indicates scales defining a factor

The Eigenvalues for Factors 1 and 2 were 13.68631 and 1.66819 respectively. The Percentages of Variance for Factor 1 was 65.2% and for Factor 2, 7.9%. The cumulative total was 73.1%.

In addition to performing reliability measures on the overall instrument, reliability analysis of the five factors established in the DeStephen and Hirokawa (1988) study yielded the results presented in Table IV.

In the DeStephen and Hirokawa study, Factor 1 reflected members' feelings of agreement satisfaction and commitment toward the groups' decision and included Items 1,2,3,4, and 5. Factor 2 consisted of Items 18, 19, 20, and 21 and measured members' feelings about their individual ability to participate in the group process. Items 10, 11, and 12 constituted Factor 3 and reflected members' feelings about group member relationships during the decision making experience. Factor 4 reflected members' feelings about the effectiveness of their individual participation in the group activities and included Items 14, 15, and 16. Finally, Factor 5 reflected the members' feelings regarding the effectiveness of decision-making techniques used in the group and included Items 6 and 9. Finally, the reliability analysis revealed a Grand Mean of 3.9917 in this study.

Table IV

Consensus Instrument Reliability: Factor Means and Alphas

		Mean	Alpha	Standardized Item Alpha
Factor 1	Group Decision	4.2087	.9685	.9688
Factor 2	Individual Opportunity	4.0543	.9183	.9247
Factor 3	Relationships	4.0326	.8943	.8988
Factor 4	Individual Effectiveness	3.9203	.9352	.9357
Factor 5	Decision Process	3.9348	.8723	.8723

Note: N = 17

Results of Analysis for RQ 2

Is there a relationship between expert ratings of team effectiveness and team ratings of effectiveness as measured by the Consensus Instrument?

Three expert raters independently grouped the 20 sample teams into 4 groups in terms of level of effectiveness, with group 1 representing the highest level of effectiveness, and group 4 representing the lowest level of effectiveness. Each group was comprised of four teams. (See Table V for Expert Ratings by Team.)

Inter-rater Reliability

Inter-rater reliability was determined by using two reliability measures. The resulting overall inter-rater reliability was found to be $r = .64$ at the $p < .05$ level. Because this measure was lower than the goal of $r \geq .7$, Holsti's (1969) formula was also employed.

Holsti's Formula

$$\text{Coefficient of Reliability} = \frac{3M}{N1+N2+N3}$$

M = the number of rating decisions upon which raters agree

N1, N2, N3 = the number of decision made by each rater

The overall C.R. = .45, employing Holsti's formula, was lower than expected. However, issues of complexity and numbers of raters raised by Holsti (1969) will be discussed in the following chapter.

Results of Analyses of Variance

Initially, it was planned to perform the analyses for the second research question using the expert raters' average groupings. However, because inter-rater reliability was low, it was decided to perform the analyses using each individual rater's rankings.

Table V**Team Effectiveness Ratings by Individual Raters**

Team Number	Rater 1	Rater 2	Rater 3
1	3	3	1
2	1	1	1
3	4	2	3
4	3	4	3
5	2	3	4
6	2	3	4
7	1	1	1
8	3	4	3
9	4	2	4
10	2	3	3
11	1	1	2
12	3	4	3
13	3	4	2
14	1	3	2
15	2	2	4
16	4	4	4
17	1	1	2
18	2	2	1
19	4	2	1
20	4	2	2

1 = Highest level of effectiveness 2 = Second highest level of effectiveness

3 = Third highest level of effectiveness 4 = Fourth highest level of effectiveness

A series of analyses of variance (ANOVAs) were conducted for each of the three expert raters by their individual groupings. Since, in the present study, only two factors emerged with the first factor accounting for 65.2% of the variance and since the overall reliability of the Consensus Instrument was $\alpha = .9729$, analysis in this thesis was conducted using the overall Consensus Instrument score. The dependent variable was the mean score for the Consensus Instrument; the independent variables were the individual groupings made by each expert rater. The emphasis of the analysis was on the overall Consensus Instrument score; however, as an additional measure, it was decided to perform one-way analyses of variance using the factor structure reported as being valid in the DeStephen and Hirokawa (1988) study.

The series of one-way analyses of variance conducted for Expert Raters 1 and 2, by their individual ratings of effectiveness for each team, yielded no significant differences among teams in terms of the overall instrument mean scores at the $p \leq .05$ level. Neither were there any significant results by factor means at the $p \leq .05$ level.

The one-way analysis of variance conducted for Expert Rater 3 also did not yield significant results for the overall instrument mean scores at the $p \leq .05$ level (See Table VI.) However, when the factor structure reported as valid by DeStephen and Hirokawa (1988) was used, the additional ANOVAs conducted on factor means for groups rated by Expert Rater 3 did yield significant results in three tests. These results of the individual tests are presented in Table VII through Table XI.

There was a negative correlation between Expert Rater 3's ratings and Group 4's mean ratings of effectiveness on Factors 2, 4, and 5. That is Group 4, the teams rated as least effective by this individual rater, reported consistently higher levels of consensus than did the other groups rated by the expert, as indicated in Table VI. Factor 2 (Items 18 through 21) measured members' feelings about their individual ability to participate.

Group 4 had significantly higher levels of consensus on this factor than did Groups 1 and 2, groups rated first and second in levels of effectiveness by this rater. On the test conducted on Factor 4 (Items 14 through 16), items reflecting members' feelings about the effectiveness of their individual participation, Group 4 had a significantly higher level of consensus than all other groups.

Finally, on Factor 5 (Items 6 and 9), concerning the effectiveness of group decisions and techniques, Group 4 reported significantly higher levels of consensus than did Group 1, the group rated highest, in terms of effectiveness, by this expert rater.

Table VI

One-way Analysis of Variance for Overall Consensus Instrument / Expert Rater 3

Groupings

<i>Source</i>	<i>df</i>	<i>MS</i>	<i>F</i>	<i>p</i>
Between Groups	3	1.2009	2.2059	.0928
Within Groups	91	.5444		
Total	94			

Table VII

One-way Analysis of Variance for Factor 1 (Individual Opportunity) / Expert Rater

3 Groupings

<i>Source</i>	<i>df</i>	<i>MS</i>	<i>F</i>	<i>p</i>
Between Groups	3	1.0969	1.3782	.2545
Within Groups	91	.7959		
Total	94			

Table VIII**One-way Analysis of Variance and Student Newman Kuels Procedure for Factor 2****(Individual Opportunity)/ Expert Rater 3 Groupings**

Source	df	MS	F	p
Between Groups	3	1.5616	2.9338	.0376
Within Groups	91	.5323		
Total	94			

Student Newman Kuels Procedure:

Groups:	Group 1	Group 2	Group 3	Group 4
Means*	3.8796 _a	3.9808 _a	3.9125	4.4432 _b

*The higher the mean, the higher the consensus level. Means with common subscripts do not differ significantly from each other. $p \leq .05$.

Table IX**One-way Analysis of Variance for Factor 3 (Relationships)/****Expert Rater 3 Groupings**

Source	df	MS	F	p
Between Groups	3	.8913	1.0254	.3852
Within Groups	91	.8692		
Total	94			

Table X

**One-way Analysis of Variance and Student Newman Kuels Procedure for Factor 4
(Individual Effectiveness)/ Expert Rater 3 Groupings**

Source	df	MS	F	p
Between Groups	3	2.3374	3.4869	.0189
Within Groups	91	.6703		
Total	94			

Student Newman Kuels Procedure:

Groups:	Group 1	Group 2	Group 3	Group 4
Means*	3.7284 _a	3.7949 _a	3.6500 _a	4.3636 _b

*The higher the mean, the higher the consensus level. Means with common subscripts do not differ significantly from each other. $p \leq .05$.

Table XI

**One-way Analysis of Variance and Student Newman Kuels Procedure for Factor 5
(Decision Process)/ Expert Rater 3 Groupings**

Source	df	MS	F	p
Between Groups	3	2.2388	2.6800	.0516
Within Groups	91	.8354		
Total	94			

Student Newman Kuels Procedure:

Groups:	Group 1	Group 2	Group 3	Group 4
Means*	3.6481 _a	3.9038	3.9250	4.3864 _b

*The higher the mean, the higher the consensus level. Means with common subscripts do not differ significantly from each other. $p \leq .05$.

Results for RQ3

The results for the third research question, "Regardless of task, are there differences in the levels of consensus based on the sociodemographic variables of gender, age, or title classification?", were mixed.

T-tests were performed to compare the means of consensus by gender; results of all comparisons are shown in Table XII. Again, the primary emphasis of the analysis was on the overall instrument score, and the first test compared mean consensus levels on the overall instrument by males and females. The obtained t-value, (-2.30, $p \leq .024$), indicated that females' overall consensus levels were significantly higher than were males' consensus levels. Again, as an additional measure, t-tests were performed using the factor structure, purported as valid, of DeStephen and Hirokawa (1988). T-tests for Factors 1 and 5 by gender also yielded significant results.

On the t-test of Factor 1, $t = -2.02$, $p \leq .048$. Factor 1 (Items 1 through 5) consisted of items which measured members' feelings of agreement, satisfaction, and commitment toward the teams' decisions. These results indicated significantly higher agreement for females. Females also reported higher consensus levels on Factor 5 than did males, $t = -3.28$, $p \leq .002$. Factor 5 tapped members' feelings regarding the effectiveness of decision-making techniques and task organization. T-test analyses for Factors 2, 3, and 4, by gender, were not statistically significant.

One-way analyses of variance, conducted to test differences in level of consensus by age, yielded no statistically significant results at the $p \leq .05$ level. Additionally, ANOVA tests yielded no statistically significant results, at the $p \leq .05$ level, for difference in mean consensus levels by title classification.

Table XII**T-Test for Mean Consensus Levels (Overall Score and Factors) by Gender**

	Subset	M	SD	T-value	p
Overall Score	Males	3.7737	.829	-2.30	.024
	Females	4.1471	.591		

DeStephen and					
Hirokawa Factors:					
Factor 1: (Group Decision)	Males	3.9900	1.119	-2.02	.048
	Females	4.3891	.656		
Factor 2: (Indiv. Opportunity)	Males	3.8813	.881	-1.74	.087
	Females	4.1636	.624		
Factor 3: (Relationships)	Males	3.8417	1.107	-1.77	.082
	Females	4.2000	.758		
Factor 4: (Indiv. Effectiveness)	Males	3.7583	.813	-1.18	.242
	Females	3.9636	.874		
Factor 5: (Decision Process)	Males	3.5750	1.083	-3.28	.002
	Females	4.2182	.712		

CHAPTER FOUR

Discussion

RQ1: Can the reported reliability and the factor structure of the DeStephen and Hirokawa Consensus Instrument be replicated?

The answer to this question, in part, is a tentative "yes." The test of Cronbach's alpha in this study was $\alpha = .9729$ which was higher than the alpha of .8906 reported by DeStephen and Hirokawa. These results support the previous study in that the internal consistency of the measure is extremely high.

On the other hand, one should be cautious in interpreting the results of the replication performed for this study. The high alpha obtained in this study is due, in part, to the fairly high levels of consensus reported by the subjects. The Grand Mean of 3.9917, obtained in this study, is less than .01 lower than "high consensus levels," as operationalized in previous research (e.g. DeStephen, 1983; DeStephen & Hirokawa, 1988.)

The factor structure obtained by DeStephen and Hirokawa (1988) was not replicated in this thesis. The DeStephen and Hirokawa study, five distinct factors accounted for a cumulative total of 60% of the variance. In this study, the principal-components factor analysis, using varimax rotation, demonstrated that two distinct factors accounted for a cumulative total of 73% of the variance (See Table III).

The failure to replicate previous findings may be due to the small sample size of this study, $N=95$, as compared to the sample size of the DeStephen and Hirokawa study, $N=234$.

Although the factors did not cluster as expected in this study, there was a clear clustering of the scales into conceptually unified factors. The first factor, as stated previously, consisted of members' feelings of agreement with the team decisions, process,

and relationships. This factor accounted for 65.2% of the variance. This finding certainly supports the conclusion (DeStephen & Hirokawa, 1988) that task is a major aspect in group consensus levels. The second factor, accounting for 7.9% of the variance, reflected members' feelings of individual effectiveness and opportunity to participate. Taken together, the results of this study support the previous findings that indicate that Consensus Instrument is an encompassing and reliable measure of consensus.

RQ2: Is there a relationship between expert rankings of team effectiveness and team ratings of effectiveness as measured by the Consensus Instrument?

As indicated earlier, the overall inter-rater reliability, $r = .64$, $p \leq .05$, was lower than expected. There may be several explanations for this result. The raters were asked to rank teams in terms of the effectiveness of their decision-making into four categories, from the highest level of effectiveness to the 4th highest level of effectiveness. Because the categories were somewhat ambiguous, the raters had to make fine distinctions between categories. When raters have to make judgements on difficult tasks, low reliability often results (Holsti, 1969). Raters may have had different frameworks from which they made their decisions on rankings, which could result in the lack of agreement in this thesis.

Additionally, only three expert raters were used. It was decided to include these three raters because of their familiarity with the quality teams and with the "Distinction Through Quality" initiative. However, this convenience sample may also have contributed to lower inter-rater reliability. In all probability, adding additional expert raters would have broadened the base of consensus among raters (Block as cited in Holsti, 1969).

For the most part, no relationship was found between expert rankings of team effectiveness and team ratings of effectiveness as measured by the Consensus Instrument. No significant differences in mean consensus levels were found when using the groupings

made by two of the expert raters. The high consensus levels reported in this study may have contributed to these results. With a Grand Mean of 3.9917, the consensus levels were skewed toward the upper end of the scales (See Table II).

However, there were significant differences in mean consensus levels for the groups of teams ranked by Expert Rater 3. This unexpected discovery was that the teams ranked least effective by Expert Rater 3, (Group 4), had higher levels of consensus than did the other groups, ranked higher by the expert, on each of the six tests. Although the results were not significant for the overall instrument score, the results were statistically significant on three tests of individual factors. Group 4 reported significantly higher consensus levels, than did all other groups, regarding their feelings about the effectiveness of their individual participation. They also had significantly higher consensus levels than two other groups did regarding members' feelings about their individual opportunity to participate. Finally, Group 4 reported significantly higher levels of consensus than did Group 1 on the factor that concerned the effectiveness of group decisions and techniques.

There are two potential explanations for the significant differences in mean consensus levels for the groups ranked by Expert Rater 3. The first explanation is simply that the differences in results are an artifact of this particular group. In this case, little more needs to be said concerning the results. Members expressed satisfaction with the teams' decisions and with their individual participation.

However, it is possible that the high consensus levels for Group 4, compared to its ranked effectiveness, could represent the phenomenon of groupthink. Janis (1972) defined *groupthink* as the extreme concurrence sought by decision-making groups. Janis's view of *groupthink* is as a process in which group members attempt to maintain a shared, positive view of the functioning of the group.

J.C. Turner (1981) suggested that groups given a social identity have a tendency to seek positive distinctiveness for the "in group" and tend to maintain a motivational bias for positive self-esteem. As a result, members develop a positive image of the group and are motivated to protect that image. Lower quality decisions are generally associated with groupthink. The structure and processes of Quality Improvement teams could contribute to groupthink symptoms. Employees are invited to become team members and are charged with coming to consensus or agreement on issues which, if resolved, could have considerable positive effects.

Although it would be a mistake to attribute the results obtained in this research question to groupthink, it remains a possible explanation. Future research on Quality Improvement Programs should investigate the potential for this phenomenon.

RQ3: Regardless of task, are there differences in the levels of consensus based on the sociodemographic variables of gender, age, or title classification?

In this thesis, there were no significant differences in levels of consensus based on the sociodemographic variables of age or title classification. This is contrary to the results of earlier studies of quality teams. In previous studies, (Fenwick & Olson, 1986; Tang et al., 1989) these variables had had positive effects on participation and satisfaction. Again, the small and unequal sample sizes for the variables of age and title classification may have contributed to these results.

On the other hand, the results of the analyses conducted for the variable of gender yielded significant differences between males' and females' levels of consensus. Females reported significantly higher mean consensus levels at the $p \leq .05$ level on the overall instrument and on members' feelings of agreement, satisfaction, and commitment toward the teams' decisions. Females also had significantly higher mean consensus levels at the

$p \leq .001$ level regarding feelings of agreement which decision-making techniques and task organization.

These results support the findings of earlier research that females, in particular, have higher aspirations for participation in decision making in the workplace (Fenwick & Olson, 1986). Because hospitals often have many more female employees than male employees, the results have implications for quality teams. Mabry's (1985) suggestions that groups need some gender mix should be heeded by those forming quality teams.

CHAPTER FIVE

Conclusions, Limitations, and Recommendations

Conclusions

The present thesis results support the previous findings by DeStephen and Hirokawa (1988) that the Consensus Instrument is a reliable and comprehensive measure of consensus. However, the utility of the instrument remains in question. DeStephen (1983) found that a consensus instrument administered at the end of a project often yielded results at the high end of the continuum. DeStephen and Hirokawa (1988) tested the instrument over time with little variation in consensus levels and suggested that the Consensus Instrument should be tested in groups with a charge and a given structure within organizations.

A primary goal of this thesis was to determine if the DeStephen and Hirokawa instrument proved to be a utilitarian measure of consensus in the organizational setting. In the present study, the results did not reflect strong variations in consensus level, even after a period of four years for some of the sample teams. Measuring consensus in consensus-seeking groups, where high levels of consensus are almost inevitable, is probably engaging in a classical "drunkard's search." With this conclusion, it becomes important to find new methods to increase our understanding of consensus.

As another goal of this thesis, external stakeholders' rankings of effectiveness were measured against team ratings of effectiveness as measured by the Consensus Instrument. Because of low inter-rater reliability, the decision was made to test individual rater's assessments against the ranked groups' mean levels of consensus on the overall instrument and individual factors. No significant differences were found in the one-way analyses of variance conducted for two of the experts' rankings. However, as an unexpected result, one raters' assessments were negatively correlated with an individual group's ratings. As

discussed earlier, the group of teams ranked lowest by one expert rater (stakeholder) reported significantly higher consensus levels than did other groups in three of the tests. Few judgments may be made about these result, since they may simply be an artifact of the group. However, future research may investigate the potential for groupthink in consensus seeking groups, such as quality teams.

Finally, this thesis sought to determine if there would be differences in consensus levels on the basis of sociodemographic variables. When age and title classification were the independent variables, there were no significant differences, which was contrary to earlier findings. However, the results of the analyses conducted for the variable of gender supports previous findings in that females in this study reported higher consensus levels than did males.

Limitations

There were several limitations in this thesis. The first was the small sample size of $N = 95$. Although the overall response rate was high, 70%, the small sample may have affected some of the results. However, while scholars often call for more research to be conducted in the field, appropriate samples may be difficult to obtain in organizations.

Subject mortality was also a limitation in this study. While 136 instruments were mailed to all present employees who had been on quality teams, 44 employees had left the organization between the time they served on a quality team and the time of this study. Turnover rates are normally high in some of the health professions. For example, nurses are often young and female, and nursing traditionally has high turnover rates. While this is common, it is possible that employees who left the organization may not have expressed high levels of consensus as did the sample for this study.

The lack of variance, resulting from high levels of consensus, was also a limitation in this thesis. Responses were skewed toward the upper end of the continuum, (possible ranges were from 1 to 5), with a grand mean of 3.9917. This lack of variance certainly resulted in the few significant differences found in the study.

Another limitation in the study was low inter-rater reliability. Pre-testing of inter-rater reliability was not conducted for this thesis. Conducting a pre-test may have resulted in changing the methodology of conducting the expert raters' assessments. For example, a common framework for assessment or a set of criteria may have been provided, or additional raters could have been obtained. The methodology used for obtaining stakeholders' rating should be tested in other research.

Finally, the Consensus Instrument used in the study may have presented a limitation. While the present research supported earlier findings that the instrument is reliable and conceptually valid, it may not be particularly useful. As discussed earlier, measuring consensus in consensus-seeking groups may be regarded as a "drunkard's search." This limitation may be solved by finding new methods for studying consensus decision making groups.

Recommendations

Recommendations for Quality Improvement Programs

Some of the results of this study have important implications for those responsible for implementing or maintaining Quality Improvement Programs. Some of the important implications involve sociodemographic characteristics of team members. QIPs are designed to involve "front-line" workers in decision-making processes. However, in this study, only 6% of the respondents reported they were Clerical/Support personnel, while 94% were faculty, administrative, managerial, and professional employees. While

hospitals certainly have a large number of professional employees, it seems that this title classification was not adequately represented in the sample. While it is possible that a high ratio of Clerical/Support personnel did not participate or had left the organization, those responsible for these programs should ensure participation by all levels of employees.

Another sociodemographic variable which should be considered in forming quality teams is the age of team members. Only six respondents reported their age as being between 20 and 29. Again, it is possible that younger employees did not participate or had left the organization. However, those forming teams should consider the inclusion of all age groups. Research does support that fact that younger employees aspire to participation in the organization (Fenwick & Olson, 1986), and their inclusion may lead to higher satisfaction.

While Quality Improvement teams at the sample organization all have trained facilitators, the results of this study have some implications for the training of these facilitators. From the review of the literature, it is apparent that the composition, charges, processes, and/or structures of quality improvement teams may make them particularly susceptible to the potential for groupthink (Turner et al., 1992). Facilitators should receive additional training in order to understand the antecedent conditions of groupthink and to minimize its potential effects.

The review of the literature also pointed out an interesting paradox which applies to quality improvement programs. Putnam and Stohl (1990) maintained that a self-destructive paradox could occur when group members must set aside external affiliations while implicitly being held accountable for them. This may be the case in many quality teams when members are asked to consider the whole or the "big picture," but are

ultimately accountable to their functional units. This paradox should, at least, be recognized as it may affect team success.

Finally, the anecdotal comments compiled in this study have important implications for quality programs. While 11% of the respondents made positive comments about the teams and/or team experiences, another 11% expressed dissatisfaction with the implementation process. In other words, while they expressed high levels of consensus on the instrument and commented positively on the team experience, they felt their teams' recommendations were either not implemented at all or were not fully implemented. Eight respondents also criticized quality improvement methods. Although these were anecdotal comments and may be the perceptions of a few respondents, the ultimate integrity and success of the quality program at the sample organization could be affected. It seems that more care should be taken with implementation and with external communications in order to improve the program. (See Appendix C for complete comments.)

Putnam and Stohl (1990) suggested that teams must be in contact with external groups to facilitate acceptability and implementation of proposals. Increasing this type of activity in quality teams might aid the ultimate success of programs. Kanter (1982) maintained that Quality Teams only provide workers the illusion of control. She stated that teams provide input which the organization subsequently ignores. In order to ensure the success of Quality Improvement, the organization and those responsible for the program must implement the recommendations fully and must communicate the implementation to the original teams. Only then will workers view participation as authentic.

Implications for Future Research

The present study sought to examine consensus decision-making in the organizational setting, and some significant results were found. However, because of

some of the limitations previously discussed and because of the general approach, there are several suggestions for future research. Putnam and Stohl (1990, pg. 251) maintain that traditional approaches fall short in capturing the dynamics, fluidity and complexity of real group experiences. Merely moving research to actual field settings may not uncover important key dimensions.

Therefore, different methodologies might be used to examine consensus decision making groups such as quality teams. Sabourin and Geist (1990) conducted discourse analytic studies in organizations. They found that decision premises, proposals, and argument patterns emanated from the external context. This would be an appropriate methodology for examining consensus decision making. Researchers could analyze the origins and patterns of development of decisions throughout the quality improvement process. Through this type of analysis, researchers could determine which team members' suggestions ultimately are adopted as decisions. Interrelation diagraphs could be employed to analyze coalition building among team members, as those coalitions relate to the decision making process. This method of careful analysis could also provide valuable information about the relative amount of participation of all team members.

Future research on quality teams should also include the assessments of external stakeholders. The methodology of the present study may be refined for future studies, or different methodologies might be employed. For example, interviews might yield richer information regarding external stakeholders' perceptions of the internal group process and decision making. Studies could also examine the boundary spanning activities of various quality improvement teams to determine if external communication and interaction has a positive effect upon decision acceptance or implementation.

Consensus decision making in organizations is more complex than the present study suggests. Descriptive, qualitative methods may more effectively capture the complex processes, constraints, dimensions, and deliberations of consensus decision making in groups in the organizational setting.

REFERENCES

- Abbott, M. L. (1987). Looking closely at quality circles: Implications for intervention. *Clinical Sociological Review*, 5, 119-131.
- Ancona, D. G. & Caldwell, D. F. (1988). Beyond task and maintenance: Defining external functions in groups. *Group and Organization Studies* 13, 468-494.
- Baker, P. M. (1988). Participation in small groups: Social, physical, and situational predictors. *Small Group Behavior* 19, 3-18.
- Bottger, P. C., & Yetton, P. W. (1987). Improving group performance by training in individual problem solving. *Journal of Applied Psychology*, 72, 651-657.
- Buch K., & Spangler, R. (1990). The effects of quality circles on performance and promotions. *Human Relations*, 43, 573-582.
- Burgoon, J. K. (1977). Unwillingness to communicate as a predictor of small group discussion behaviors and evaluations. *Central States Speech Journal*, 28, 122-133.
- Cragan, J. D., & Wright, D. W. (1990). Small group research of the 1980s: A synthesis and critique. *Communication Studies*, 41, 212-236.
- Davis, J. H. (1992). Some compelling intuitions about group consensus decision, theoretical and empirical research, and interpersonal phenomena: Selected examples, 1950-1990. *Organizational Behavior and Human Decision Processes*, 52, 3-38.
- Davis, L. V., & Toseland, R. W. (1987). Group versus individual decision-making: An experimental analysis. *Administration in Social Work*, 11, 49-58.
- DeStephen, R. S. (1983) Group interaction differences between high and low consensus groups. *Western Journal of Speech Communication*, 47, 340-363.

- DeStephen, R. S., & Hirokawa, R. Y. (1988). Small group consensus: Stability of group support of the decision, task process, and group relationships. *Small Group Behavior, 19*, 227-239.
- Donovan, M., & Jury, P. (1983). The impact of quality circles on aspects of the work climate. Transactions of the 5th Annual Conference of the International Association of Quality Circles, (pp. 463-470).
- Fenwick R., & Olson, J. (1986). Support for worker participation: Attitudes among union and nonunion employees. *American Sociological Review, 51*, 505-522.
- Ferris, G. R., & Wagner, J. A. (1985). Quality circles in the United States: A conceptual reevaluation. *Journal of Applied Psychology, 71*, 155-167.
- Ganster, D. C.; Williams, S.; & Poppler, P. (1991). Does training in problem-solving improve the quality of group decisions? *Journal of Applied Psychology, 76*, 479-483.
- Geist, P., & Chandler, T. (1984). Account analysis of influence in group decision-making. *Communication Monographs, 51*, 67-78/
- Gersick, C.J.G. (1988). Time and transition in work teams: Toward a new model of group development. *Academy of Management Journal, 31*, 9-41.
- Gladstein, D. L. (1984). Groups in context: A model of task group effectiveness. *Administrative Science Quarterly, 29*, 499-517.
- Gouran, D. S. (1984). The paradigm of unfulfilled promise: A critical examination of the history of research on small groups in speech communication. In T. W. Benson (Ed.), *Speech communication in the twentieth century*. Carbondale, IL: Southern Illinois University Press.

- Hackman, J. R., & Morris, C. G. (1975). Group tasks, group interaction process, and group performance effectiveness: A review and partial integration. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 8, pp. 47-99). San Diego, CA: California Academic Press.
- Hackman, J. R., & Oldham, G. R. (1980). Motivation through the design of work: A test of a theory. *Organizational Behavior and Human Performance*, 16, 250-279.
- Hill, M. (1982). Group versus individual performance: Are N + 1 heads better than one? *Psychological Bulletin*, 91, 517-539.
- Hirokawa, R. Y. (1980). A comparative analysis of communication patterns within effective and ineffective decision-making groups. *Communication Monographs*, 47, 312-321.
- Hirokawa, R. Y., & Johnston, D. D. (1989). Toward a general theory of group decision making: Development of an integrated model. *Small Group Behavior*, 20, 500-523.
- Holsti, O. R. (1969). *Content Analysis for the Social Sciences and Humanities*. Reading, MA: Addison-Wesley Publishing Co.
- Hrycenko, I., & Minton, H. L. (1974). Internal-external control, power position and satisfaction in task-oriented groups. *Journal of Personality and Social Psychology*, 30, 871-878.
- Janis, I. L. (1972). *Victims of groupthink*. Boston: Houghton-Mifflin.
- Kanter, R. M. (1982). Dilemmas of managing participation. *Organizational Dynamics*, 11, 5-27.
- Landon, D., & Mouton, S. A. (1986). Quality circles: What's in them for employees? *Personnel Journal*, 65, 23-27.

- Lawler, E. E. III, & Mohrmann, S. A. (1987). Quality circles: After the honeymoon. *Organizational Dynamics*, 15, 42-54.
- Littlejohn, S. W. (1989). *Theories of human communication (3rd Ed.)*, Belmont, CA: Wadsworth.
- Liverpool, P. R. (1990). Employee participation in decision-making: An analysis of the perceptions of members and nonmembers of quality circles. *Journal of Business and Psychology*, 4, 411-422.
- Mabry, E. A. (1985). The effects of gender composition and task structure on small group interaction. *Small Group Behavior*, 16, 75-96.
- McGrath, J. E., & Kravitz, D. A. (1982). Group research. *Annual Review of Psychology*, 33, 195-230.
- Michaelsen, L. K.; Watson, W. E.; & Black R. H. (1989). A realistic test of individual versus group decision-making. *Journal of Applied Psychology*, 74, 834-837.
- Miner, F. C. (1984). Group versus individual decision-making: An investigation of performance measures, decision strategies, and process losses/gains. *Organizational Behavior and Human Performance*, 33, 112-124.
- Mohr, W. L., & Mohr, H. (1983). *Quality Circles: Changing Images of People at Work*. Reading, MASS: Addison-Wesley.
- Park, S. J. (1991). Estimating the success rates of quality circle programs: Public and private experiences. *Personnel Administration Quarterly*, Spring, 133-146.
- Paulus, P. B., & Dzindolet, M. T. (1993). Social influence processes in group problem-solving. *Journal of Personality and Social Psychology*, 64, 575-586.
- Paulus, P. B.; Dzindolet, M. T.; Poletes, G; & Camancho, L. M. (1993). Perceptions of performance in group brainstorming: The illusion of group productivity. *Personality and Social Psychology Bulletin*, 19, 78-89.

- Pendell, S. D. (1990). Deviance and conflict in small group decision-making: An exploratory study. *Small Group Research, 21*, 393-403.
- Putnam, L. L., & Stohl, C. (1990) Bona fide groups: A reconceptualization of groups in context. *Communication Studies, 41*, 248-265.
- Russell, R. (1988) Forms and extent of employee participation in the contemporary United States. *Work and Occupations, 15*, 374-395.
- Sabourin, T. C., & Geist, P. (1990). Collaborative production of proposals in group decision-making. *Small Group Research, 21*, 404-427.
- Schwenk, C.R. & Cosier, R. A. (1992). Effects of consensus and devil's advocacy on strategic decision-making. *Journal of Applied Social Psychology, 22*, 126-139.
- Schonberger, R. J. (1994). Human resource management lessons from a decade of total quality management and reengineering. *California Management Review, (Summer, 1994)*: 109-123.
- Smeltzer, L. R., & Kedia, B. (1985). Knowing the ropes: Organizational requirements for quality circles. *Business Horizons, 28*, 30-34.
- Steel, R. P.; Mento, A. J.; Dilla, B. L.; Ovalla, N. K; & Lloyd, R. F. (1985). Factors influencing the success and failure of two quality circles. *Journal of Management, 11*, 99-119.
- Steel, R. P., & Shane, G. S. (1984). Evaluation research on quality circles: Technical and analytical implications. *Human Relations, 39*, 449-468.
- Steiner, I. D. (1972). *Group process and productivity*. New York: Academic Press.
- Stohl, C. (1986). Quality circles and the changing patterns of communication. In M. L. McLaughlin (Ed.), *Communication Yearbook 9 (pp. 511-531)*. Beverly Hills: Sage.

- Stohl, C., & Schell, S. E. (1991). A communication based model of a small group dysfunction. *Management Communication Quarterly*, 5, 90-110.
- Tang, T. L. P., Tollison, P. S., & Whiteside, H. D. (1989). Quality circle productivity as related to upper management attendance. *Journal of Management*, 15, 101-113.
- Thompson, P. C. (1982). *Quality circles: How to make them work in America*. New York: AMACOM.
- Tortorich, R.; Thompson, P.; Orfan, C; Layfield, D; Dreyfuss, C; & Kelley, M. (1981). Measuring organizational impact of quality circles. *The Quality Circle Journal*, 4, 24-34.
- Townsend, P. L., & Gebhardt, J. E. (1990). *Commit to quality*. New York: Wiley.
- Tuckman, A. (1994). The yellow brick road: Total quality management and the restructuring of organizational culture. *Organization Studies*, 15, 727-751.
- Turner, J. C., (1981). The experimental social psychology of intergroup behavior. In J. C. Turner and H. Giles (Eds.), *Intergroup Behavior* (pp. 66-101) Chicago: University of Chicago Press.
- Turner, M. E.; Pratkanis, A. R.; Probasco, P.; & Leve, C. (1992). Threat, cohesion, and group effectiveness: Testing a social identity maintenance perspective on groupthink. *Journal of Personality and Social Psychology*, 63, 781-796.
- Vandervelde, M. (1979). The semantics of participation. *Administration in Social Work*, 3, 65-77.
- Van de Ven, A. H., & Delbecq, A. L. (1974). The effectiveness of nominal, delphi, and interaction group decision-making processes. *Academy of Management Journal*, 17, 605-621.
- Wall, V. D.; Galanes, G. J.; & Love, S. B. (1987). Small, task-oriented groups: Conflict, conflict management, satisfaction and decision quality. *Small Group Behavior*, 18, 31-55.

- Wall, V. D., & Nolan, L. L. (1986). Perceptions of inequity, satisfaction, and conflict in task-oriented groups. *Human Relations, 39*, 1033-1052.
- Watson, W.; Michaelsen, L. K.; & Sharp, W. (1991). Member competence, group interactions and group decision-making: A longitudinal study. *Journal of Applied Psychology, 76*, 803-809.
- Wendt, R. F. (1994). Learning to "walk the talk" A critical tale of the micropolitics at a total quality university. *Management Communication Quarterly, 8*, 5-45.
- Yetton, P. W., & Bottger, P. C. (1982). Individual versus group problem-solving: An empirical test of a best-member strategy. *Organizational Behavior and Human Performance, 29*, 307-321.
- Zaleznik, A., & Moment, D. (1964). *The Dynamics of Interpersonal Behavior*. New York: John Wiley.

Appendix A
IRB Exemption



Institutional Review Board
For the Protection of
Human Subjects

University of Nebraska Medical Center
Eppley Science Hall 3018
600 South 42nd Street
Box 986810
Omaha, NE 68198-6810
(402) 559-6463
Fax (402) 559-7845

March 29, 1996

Ginger Riffel
Department of Communication
UNO - 5710

IRB #: 127-96-EX

TITLE OF PROTOCOL: Consensus Decision Making of Quality Improvement Teams:
A Descriptive Study

Dear Ms. Riffel:

The IRB has reviewed your Exemption Form for the above-titled research project. According to the information provided, this project is exempt under 45 CFR 46:101b, category 2. You are therefore authorized to begin the research.

It is understood this project will be conducted in full accordance with all applicable sections of the IRB Guidelines. It is also understood that the IRB will be immediately notified of any proposed changes that may affect the exempt status of your research project.

Sincerely,

A handwritten signature in cursive script, appearing to read "Ernest D. Prentice".

Ernest D. Prentice, PhD
Vice Chairman, IRB

EDP:jlg

Appendix B
Cover Letter
and
Consensus Instrument

April 3, 1996

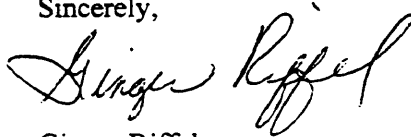
Dear Quality Improvement Team Member:

As a graduate student at the University of Nebraska at Omaha, I am doing research on group decision making for my Master's Thesis. In particular, I'm interested in the decision making of quality improvement teams.

You were a member of the quality improvement team indicated on the enclosed survey. I am asking for your help in completing my research project. Your response is important because it will provide valuable information for the "Distinction Through Quality" initiative, and it will help me complete the requirements for graduation. Of course, your participation in this project is voluntary. If you are willing, please complete the enclosed survey according to your feelings at the time your team *completed* its work, and recommendations were ready for approval by the quality council. All individual responses will be completely confidential. The sociodemographic information will be used solely for grouping responses to make comparisons.

Please complete the survey and return it in the envelope provided within the next 10 days to Ginger Riffel in Organizational Learning and Development at zip 5710. If you have any questions, please feel free to call me at extension 96358. Thanks so much for your help.

Sincerely,

A handwritten signature in cursive script that reads "Ginger Riffel". The signature is written in black ink and is positioned above the typed name.

Ginger Riffel

Decision Making Survey

Thank you for your assistance in helping me complete my research project. Please indicate how accurately the following statements describe your feelings about your experience on the quality team named below. Please answer the statements which best describe your feelings at the time your team's work was completed, and recommendations were ready to be approved by the Quality Council.

Name of Team _____

Please circle the number that tells best you felt at the time your team's work was completed. Please complete all items.

- 1 STRONGLY DISAGREE. This was almost never the way things were
2. DISAGREE. This was not usually the way things were.
3. IN BETWEEN. This sometimes was and sometimes wasn't the way things were.
4. AGREE. This was usually the way things were.
5. STRONGLY AGREE. This was almost always the way things were.

	Strongly Disagree				Strongly Agree
1. The team reached the right decisions.	1	2	3	4	5
2. I believe that our team's decisions/solutions were appropriate.	1	2	3	4	5
3. I supported the final team decisions.	1	2	3	4	5
4. I believe we selected the best alternatives available.	1	2	3	4	5
5. I would be willing to put my best effort into carrying out the team's final decisions.	1	2	3	4	5
6. I believe we approached our task in an organized manner.	1	2	3	4	5
7. This team accomplished what it set out to do.	1	2	3	4	5
8. Our team worked well together.	1	2	3	4	5
9. This team used effective decision-making techniques.	1	2	3	4	5
10. This team was a place where people could feel comfortable expressing themselves.	1	2	3	4	5
11. I liked the members of my team.	1	2	3	4	5

		Strongly Disagree				Strongly Agree
12.	I would like to work with members of my team on another similar project	1	2	3	4	5
13.	We were a closely knit team.	1	2	3	4	5
14.	I believe I contributed important ideas during the decision-making process.	1	2	3	4	5
15.	I believe I had a lot of influence on the team's decision-making.	1	2	3	4	5
16.	I contributed important information during the team's decision-making process.	1	2	3	4	5
17.	Without my input and suggestions, the team would not have come up with good decisions/solutions.	1	2	3	4	5
18.	During team meetings, I got to participate whenever I wanted to.	1	2	3	4	5
19.	I believe that the other members of the team liked me.	1	2	3	4	5
20.	Other team members really listened to what I had to say.	1	2	3	4	5
21.	I felt that I was a genuine member of the team.	1	2	3	4	5

Please complete the following by placing an "X" on the appropriate line. This information will be used only for grouping responses to make comparisons.

Gender:

____ Male

____ Female

Age:

____ 20-29

____ 30-39

____ 40-49

____ 50-59

____ 60-69

Title Classification:

____ Faculty/Administrative

____ Managerial/Professional

____ Clerical/ Support

Comments: _____

Appendix C
Comments

Figure 1

Anecdotal Comments Listed by Team

- 1: *We actually implemented 50%+ *In comparison to other, more recent "teams" or "focus groups" I have been involved with, the _____ team was a great group which accomplished its objectives. Thanks.
- 2: *I was the facilitator - appropriate to use me?
- 3: *Too bad nothing came of it.
- 4: *Project was not crafted to measure outcome measures. I doubt it had any significant long-term impact. Analysis of data was not rigorous or statistically sound.
- 7: *The project was excessively labor intensive -- much more expensive than what the project probably warranted -- As usual, secretarial support was a huge unmet need. *The DTQ team did good work but, in my opinion, it was a very inefficient way to get the work done. *Had good outcomes, almost all recommendations were put into action.
- 8: *I like working with the girls. I got to meet & know them I enjoy it.
- 9: *Was facilitator for this team. *We real had a great team. Very focused and very interested in helping the _____ patients. *We were a cohesive team. We developed a practical solution to resolve delays in chem delivery.
- 10: *Not sure my input is appropriate since I was facilitator. *Good Luck with your project! *As with all CQI teams I have observed things went well until implementation - which never occurs...
- 12: *Didn't implement anything! *The most discouraging part of this process is that none of our suggestions have been implemented. *Our team finished 11/93 - it was hard to answer some of these questions because it was so long ago. *None of the final conclusions that the team put together were ever carried out -- So this was a **big** waste of time!
- 13: *Unfortunately very few of our ideas were implemented the 1st time around. We had M.D. rep on our team who attended the majority of our meetings which was great!!

- 14: * Actual use of the COSTAR system has been minimal for a variety of reasons despite the ability of this system to completely resolve the issues behind the formation of the team. The time commitment to this process is very difficult for physicians with ever increasing demands in other non administrative areas. I would avoid future projects.
- 15: *Facilitator of team. *We never accomplished what we set out to do, because the project was never implemented. *I liked being on/the experience of a DTQ team. We worked hard to accomplish our goals. The presentation of our conclusions to the Hospital Board was the hardest thing. It was frustrating for me, however, because I never heard if our recommendations were/are implemented or not. It has been a few years. They told us we would be involved/informed about the implementation process also. It was kind of like our "baby." I never heard a word.
- 17: *Very positive experience. I would do it again if asked. Liked the process; however, sometimes wondered if actual results are obtained, carried out, and then evaluated related to the entire QMT process. *The _____ Team set out to complete its task, and all worked well together to accomplish our goals --This was a much better outcome than the more recent _____ team I served on. The second team worked well together also, but met too many obstacles in attempting to reach our goals. *Q#12 - I would no longer be an appropriate member as my job responsibilities have changed to the point were I don't have the same and necessary info. for the team to move forward. This was a high-functioning team & we had a lot of fun. To the best of my knowledge none of the recommendations have been implemented. Makes me feel like it was wasted time as no the information is over a year old.
- 18: *I enjoyed working with team - and actually anticipated our meeting days. *I truly believe we did good work but the implementation team kind of dies and not too much has been done.

- 19: *I felt the team process got bogged down in CQI techniques and took 12 months to do the obvious - and then did not follow through to be sure the changes worked. From my point of view things are not much different than before the project was started. *My QIT had many breaks in meetings during the process so the momentum and focus was hard to maintain. The time between the completion of the process and presentation also was lengthy so the enthusiasm and excitement of the project was missing at the presentation.
- 20: *You never asked if I thought it was the most efficient & best way to approach the problem -- I didn't that is also why I marked question #9 low. I believe most of the conclusions were evident before the committee met & a plan could have been reached more quickly by fewer people meeting less often. *Good Luck.

* Indicated the beginning of an individual's comments.

_____ indicates identification of specific team.